The Effect of Human Genetic Factors on Childhood Malaria in Asembo, Western Kenya

Thesis

The effect of human genetic factors on childhood malaria in Asembo, Western Kenya.

Oscar Asanya Nyangiri, BSc.

The Open University

Thesis Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

May 2015

Affiliated Research Centre

KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya

Collaborating establishments

KEMRI/Centers for Disease Control and Prevention, Kisumu, Kenya

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

Human genetic factors confer protection or susceptibility to malaria. Classical examples include sickle cell trait, which confers up to 90% protection from severe *Plasmodium falciparum* malaria and Duffy antigen negativity, which offers almost complete resistance to *Plasmodium vivax* infection. Unfortunately, such genetic factors are insufficiently understood to design interventions. I conducted the current study in Asembo, western Kenya by recruiting a paediatric cohort, characterizing malaria epidemiology, genotyping for single nucleotide polymorphisms (SNPs) to identify genetic factors affecting malaria incidence and investigating their possible relationships with naturally acquired immunity. In addition, I investigated the effect of co-inheritance of genetic factors on malaria incidence. Children under 12 years were followed up over a 6 year period (2008-2013) to calculate malaria incidence. Odds ratios for malaria were also calculated for categorical risk factors identified during a nested cross sectional survey. The parasite density threshold associated with a malaria fever was investigated through a logistic regression approach. Interactions between the gene variants, and between the gene variants and age were investigated using the likelihood ratio test (LRT). Evidence for differences in age-specific rates of malaria was also investigated by LRT. Through these studies, risk factors of malaria were identified. I established differences in age and genotype specific incidence, suggesting that malaria candidate genes affect naturally acquired immunity to malaria. In addition, this study identified pairs of genes that may interact to affect malaria incidence. Gene frequencies of various malaria candidate genes as typed in this study may be an important consideration in interpreting results of intervention trials in this population.

The high frequency of the G6PD deficiency genotypes determined in this study is clinically relevant as it determines the use of primaquine, a key drug which causes haemolysis in deficient individuals, but which is useful in malaria elimination due to its gametocytocidal activity.

DECLARATION

The work presented here is my original work and has not been presented before for the award of any degree or publication. Part of the data analysed here was from the KEMRI/CDC Health and Demographic Surveillance system and the population based infectious disease surveillance system. I wrote the proposal, sought scientific and ethical permissions, set up the genetic cohort, conducted the cross-sectional survey, assisted with genotyping, conducted data analysis and wrote the thesis. Haemoglobin electrophoresis was done with the help of colleagues at the KEMRI/CDC Programme in Kisumu. Multiplex typing was done in collaboration with the Wellcome Trust Centre for Human Genetics in Oxford, United Kingdom. Typing for α-thalassaemia was done with the help of colleagues at the KEMRI/Wellcome Trust Research Programme, Kilifi.

ACKNOWLEDGEMENTS

This work would not have been possible without the support of a great team of mentors, advisors, and partners. I owe a debt of gratitude to my supervisors, Prof. Thomas N Williams, Dr. Joel Montgomery and Dr. Simon Kariuki, for believing in me and offering inspiration, guidance, and encouragement throughout the course of my studies. I acknowledge the support of the KWTRP training department, especially Dr. Sam Kinyanjui and Liz Murabu. I am very grateful to Dr. Deron Burton and Mr. Godfrey Bigogo whose guidance was invaluable in setting up this cohort study within the disease surveillance framework, and Dr. Frank Odhiambo for facilitating access to the Health and Demographic Surveillance platform. I am indebted to study participants and their parents for making this study a reality. Philgona Otiende and Hezekiah Odhiambo, the study nurses contributed immensely to sample collection. Thanks are due to the Asembo field workers led by George and Ratwar; and Lwak hospital staff led by Dr. Ochieng' and Sister Jane who facilitated recruitment of participants. At KEMRI/CDC I acknowledge support from the malaria branch, including Stephanie, Meghna, Vincent and Peter, administrative support from Ken, Rhoda and Booker and support from laboratory colleagues especially Jim, Telesphorus and Vincent. The data management team at KEMRI/CDC led by Allan Audi and Billy Ogwel and at KWTRP led by Dr. Greg Fegan, Gideon, Mike, Amos, Mark and Ojal gave valuable insights. At KWTRP, I acknowledge the help of Alex in the library and my laboratory colleagues especially Alex, Johnstone, Adan, Metrine, Herbert, Carolyne, Sophie, Mary, Prophet and Emily and my office colleagues Juliet, Ngure, Donald, Olive and Esther who were very supportive. To my family and friends especially Maureen, Justus and Eric, thank you for your love and support. Finally, I thank God for giving me life and the opportunity to make a contribution through this study.

DEDICATION

To my parents, Agnes Kemunto Nyangiri and the late Enock Okioga Nyangiri.

Table of contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ABO	Blood grouping system which uses red cell surface antigens A, B and O
ACT	Artemisinin Combination Therapy
AL	Artemether/Lumefantrine
ALRI	Acute Lower Respiratory Illness
AMA-1	Apical Membrane Antigen- 1
An. funestus	*Anopheles funestus*
An. gambiae	*Anopheles gambiae*
An. arabiensis	*Anopheles arabiensis*
ARI	Acute Respiratory Illness
ATP2B4	ATPase, Ca++ transporting plasma membrane 4
AVPU scale	Alert, Voice, Pain, Unconscious
CAB	Community Advisory Board
cAMP	cyclic Adenosine Monophosphate
CDC	Centers for Disease Control and Prevention
CGMRC	Centre for Geographical Medicine and Research
CI	Confidence Interval
CM	Cerebral Malaria
CR1	Complement Receptor 1
Cyfu	Child Years of follow up
DBS	Dry blood spot
DNA	Deoxyribonucleic acid

DVS	Dominant Vector Species
EBL	Erythrocyte Binding Like
EIR	Entomological Inoculation Rate
ERC	Ethics Review committee
G6PD	Glucose 6 phosphate Dehydrogenase
gDNA	Genomic DNA
GPI	Glycosylphosphatidylinositol
Hb	Haemoglobin
HbAA	Normal Haemoglobin
HbAS	Haemoglobin AS - sickle cell trait
HbSS	Sickle cell anaemia
HbC	Haemoglobin C
HbF	Foetal haemoglobin
HbS	Haemoglobin S
HDSS	Health and Demographic Surveillance System
HLA	Human Leukocyte Antigen
HP	Haptoglobin 2-2
HRP-2	Histidine rich protein 2
IBD	Inherited Blood Disorders
IEIP	International Emerging Infections Program
IgE	Immunoglobulin E (Ig E)
IPTp	Intermittent Preventive Treatment of Malaria in pregnancy
ITN	Insecticide Treated Net

IRR	Incidence Rate Ratio
IRS	Indoor Residual Spraying
KEMRI	Kenya Medical Research Institute
KWTRP	KEMRI Wellcome Trust Research Programme
MAF	Malaria Attributable Fraction of fever
MARVELD3	Myelin and lymphocyte and related proteins for vesicle trafficking and membrane link, domain containing protein
3MCA	Multiple Correspondence Analysis
McC	McCoy genotypes of CR1 protein
OR	Odds Ratio
P. malariae	*Plasmodium malariae*
P. berghei	*Plasmodium berghei*
P. falciparum	*Plasmodium falciparum*
P. knowlesi	*Plasmodium knowlesi*
P. ovale	*Plasmodium ovale*
P. vivax	*Plasmodium vivax*
PBIDS	Population Based Infectious Disease Surveillance
PCR	Polymerase Chain Reaction
PEP	Primer Extension Pre-amplification
PfAPI	*Plasmodium falciparum* Annual Parasite Index
PfPR	*Plasmodium falciparum* Parasite Rate
PfRH	*P.falciparum* Reticulocyte Binding like Protein Homologue
PR	Parasite Rate

PvDBP	*P. vivax* Duffy Binding protein
RBCs	Red Blood Cells
RDTs	Rapid Detection Tests
RON-2	Rhoptry Neck Protein-2
rsnumber	reference SNP number; used as a unique identifier for SNPs
RTS, S	Repeat region of central Circumsporozoite Protein (CSP), T-cell epitope of *P. falciparum CSP,* Hepatitis B Surface antigen, S fusion protein
SES	Social Economic Status
SI	Swain Langley genotypes of CR1 protein
SMA	Severe Malarial Anaemia
SNP	Single Nucleotide Polymorphism
SP	Sulphadoxine/Pyrimethamine
SR	Sporozoite Rate
SSC	Scientific Steering Committee
NAI	Naturally Acquired Immunity
TVM	Tubular Vesicular Membranes
URTI	Upper respiratory Tract Infection
USD	US Dollar
WHO	World Health Organization
WTCHG	Wellcome Trust Centre for Human Genetics
XSS	Cross-Sectional Survey

CHAPTER 1: Background and literature review

1.1 The global malaria problem

Malaria is estimated to have caused 207 million cases globally in 2012, resulting in approximately 600,000 deaths (WHO, 2013). Although higher estimates of mortality exist depending on the measurement method (Murray *et al.*, 2012), taken together these estimates show that malaria mortality remains unacceptably high. The current high mortality estimates persist despite a 42% decrease in reported mortality between 2002 and 2012, underlining the magnitude of the global malaria problem, such that in 2012, 104 countries were still considered malaria-endemic (WHO, 2013). While sub-Saharan Africa accounts for about 80% of malaria cases, other major malaria endemic countries include parts of South America, South- and South-East Asia and the Pacific (Gething *et al.*, 2010; WHO, 2013). This burden of malaria is not uniform across age groups, with children under 5 years accounting for about 80% of malaria cases worldwide, and up to 90% of malaria mortality occurring in sub-Saharan Africa (WHO, 2013). Currently approximately 48% of the global population live in areas at risk of malaria (Hay *et al.*, 2004). Challenges still remain in malaria control, gaps in financing malaria control (WHO, 2013) and the threat of insecticide and drug resistance (Feachem *et al.*, 2010; McCann *et al.*, 2014) being among the most serious.

1.2 Malaria Biology

Five *Plasmodium* species have been shown to cause malaria in humans, *P. falciparum*, *P. vivax, P. ovale*, *P. malariae* and, most recently, *P. knowlesi* (Singh *et al.*, 2004).

Although these species differ in morphology and clinical presentation, in large part they share a common life cycle as described below.

1.2.1 The Parasite

1.2.1.1 Life cycle

Modern malaria study is anchored on the work of Laveran, Manson, Ross, and Grassi, which established the life cycle of malaria parasites in the late 19th century (Grassi, 1900; Grassi *et al.*, 1899; Laveran, 1881, 1884; Manson, 1879; R. Ross, 1897). Laveran exclusively observed protozoa in blood from individuals who presented with malarial fevers (Laveran, 1881, 1884) (reviewed by Bruce-Chwatt (Bruce-Chwatt, 1981)). Following from clues offered by evidence implicating mosquitoes in transmission of the filarial worm *Wuchereria bancrofti* by Manson (Manson, 1879), Ross showed that malaria parasites could infect particular species of mosquitoes (R. Ross, 1897), thus implicating mosquitoes as the vectors of malaria. Grassi and colleagues later showed mosquito transmission of human malaria, and noted that only female anopheles mosquitoes transmit malaria (Grassi, 1900; Grassi *et al.*, 1899).

The current understanding of the life cycle of malaria parasites in man is summarized in Figure 1.1. The life cycle can broadly be classified into asexual stages in the human host (exo-erythrocytic stage, erythrocytic stage), gametocytogenesis and sexual stages in the mosquito.

1.2.1.1.1 Asexual stage

1.2.1.1.1.1 Exo-erythrocytic stage (hepatic schizogony)

When infected mosquitoes take a human blood meal, they inject their saliva into subcutaneous tissues prompting a vasodilatory response and facilitating feeding. Sporozoites in the mosquitoes' saliva are incidentally injected into the host, migrate into the bloodstream and infect the human host. Sporozoites travel in the blood stream to the liver. The process of liver infection has been observed by fluorescent microscopy in a mouse model (Frevert *et al.*, 2005), where *P. berghei* binds to the sinusoidal cell layer, crosses the Kuppfer cells and finally binds to the hepatocyte. Cell passage through Kuppfer cells is crucial to hepatocyte invasion (Ishino *et al.*, 2004). In the liver, each sporozoite infects the liver cells (hepatocytes), asexually reproduces and develops into merozoites. Merozoites eventually rupture from the liver cells and enter into the blood stream.

1.2.1.1.1.2 Erythrocytic stage

Merozoites released from the liver rapidly invade erythrocytes, a process that has been studied by video and electron microscopy (Aikawa *et al.*, 1981; Dvorak *et al.*, 1975). Merozoites invade the erythrocytes in a series of four steps. First, contact with the red blood cell occurs at any point of the ovoid shaped merozoite. The merozoite then reorients itself so that its apical membrane is in direct contact with the erythrocyte. Secretory organelles on the apical end mediate the formation of an invaginated moving-junction with the erythrocyte membrane. The merozoite then enters the red blood cell, in the process forming a parasitophorous vacuole that surrounds the merozoite inside the erythrocyte. This complex process takes less than 2 minutes (Gilson & Crabb, 2009). Various parasite proteins mediate this process. Merozoite

Surface Proteins mediate initial contact with the erythrocyte. The *P. falciparum* erythrocyte binding like (EBL) antigen family, related to the *P. vivax* Duffy binding protein (PvDBP) is thought to play a key but redundant role in invasion. However, in *P. vivax* the role of the PvDBP is more crucial, leading to the almost complete refraction of the erythrocyte to invasion in the absence of its erythrocyte receptor, the Duffy antigen (Miller *et al.*, 1976). The *P. falciparum* reticulocyte-binding-like protein homologue (PfRH) family also mediate binding to reticulocytes and play an important role in host cell recognition and invasion. Recent interest in this family has been centred on PfRH5, which binds to basigin in the host (Baum *et al.*, 2009). Polymorphisms in this ligand have been shown to confer virulence to previously non-virulent strains of *P. falciparum*, allowing them to infect Aotus monkeys (Hayton *et al.*, 2008). A further important parasite protein identified to play a role in invasion is Apical Membrane Antigen-1 (AMA-1), which atypically binds to a parasite protein the Rhoptry Neck Protein-2 (RON-2) which is in turn bound to the red blood cell.

After invasion, the parasites grow inside the erythrocyte. This necessitates uptake of nutrients from the extracellular environment, which is thought to take place either through increased permeability of the erythrocyte membrane or through formation of channels to the extracellular environment. Kirk and colleagues (Kirk, 2001) have reviewed the evidence for increased permeability of the erythrocyte membrane following infection by merozoites and have posited that new permeation pathways of the infected erythrocyte allow uptake of nutrients and solutes. The parasites are also believed to form tubular vesicular membranes (TVM) which are channels linking them to the extracellular environment for uptake of nutrients (Lauer *et al.*, 1997). Nutrient uptake allows growth and further asexual reproduction within the red blood cell. The growth stages include ring stages and eventually schizont stages within the

erythrocyte. Within 48 hours of invasion, the erythrocyte ruptures, releasing more merozoites that in turn invade more red blood cells. The rupture of the red blood cells correlates with the symptoms of malaria experienced by the host including fever, nausea and vomiting.

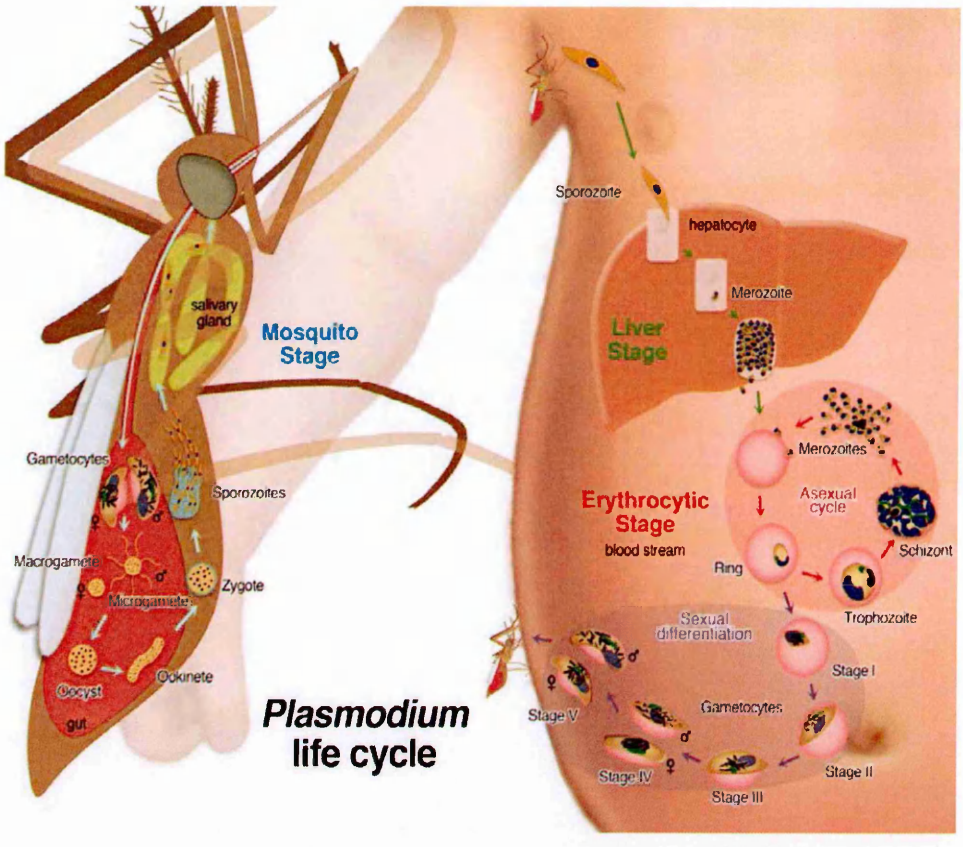
1.2.1.1.1.3 Gametocytogenesis

A subset of parasites undergo differentiation to form gametocytes, named micro-gametes (male) and macro-gametes (female). The mechanism involved is not fully understood. Recently, the parasite protein *P. falciparum* Histone Deacetylase-2 has been shown to regulate gametocyte conversion (Coleman *et al.*, 2014). Only mature gametocytes are seen in the host peripheral circulation leading to the hypothesis that immature stages sequester in deep vascular tissues. Human autopsy studies show disproportionately higher burdens of gametocytes in the tissues, especially the bone marrow (Joice *et al.*, 2014). This may happen for at least two reasons, either tissue specificity is required for gametocyte formation, or particular conditions that are found in the tissues favour the development of gametocytes. Evidence for both exists; for the latter, Kaushal and colleagues (Kaushal *et al.*, 1980) used *in-vitro* culture techniques to show that static cultures with high parasite densities favour gametocytaemia whereas gametocytes form at lower rates in conditions favouring rapid proliferation. Extracellular cyclic AMP has also been shown to induce gametocyte formation. Regarding tissue specificity, Joice and colleagues (Joice *et al.*, 2014) have shown that immature stages of gametocytes accumulate in the erythroid precursor cells, implicating these cells as specific sites for gametocytaemia (Joice *et al.*, 2014).

1.2.1.1.2 Sexual cycle (sporogony)

During a blood meal, female anopheles mosquitoes ingest gametocytes from an infected host. Macrogametes and microgametes fuse to form zygotes that mature to form motile ookinetes, which attach to the mosquito's midgut. Eventually, ookinetes pass through the midgut and attach to the other side of the midgut surface where they undergo meiosis (Sinden & Hartley, 1985). Mitotic division results in formation of mature oocysts, which contain haploid sporozoites. Oocysts then rupture, allowing the sporozoites to access the coelomic cavity and migrate to the salivary glands. When the mosquito feeds on human hosts, the parasites in the saliva are injected into the human host, thus completing the cycle of infection.

Figure 1. 1 Life cycle of malaria parasites



Adapted from http://ucrtoday.ucr.edu/19520/falciparum-life-cycle-final.

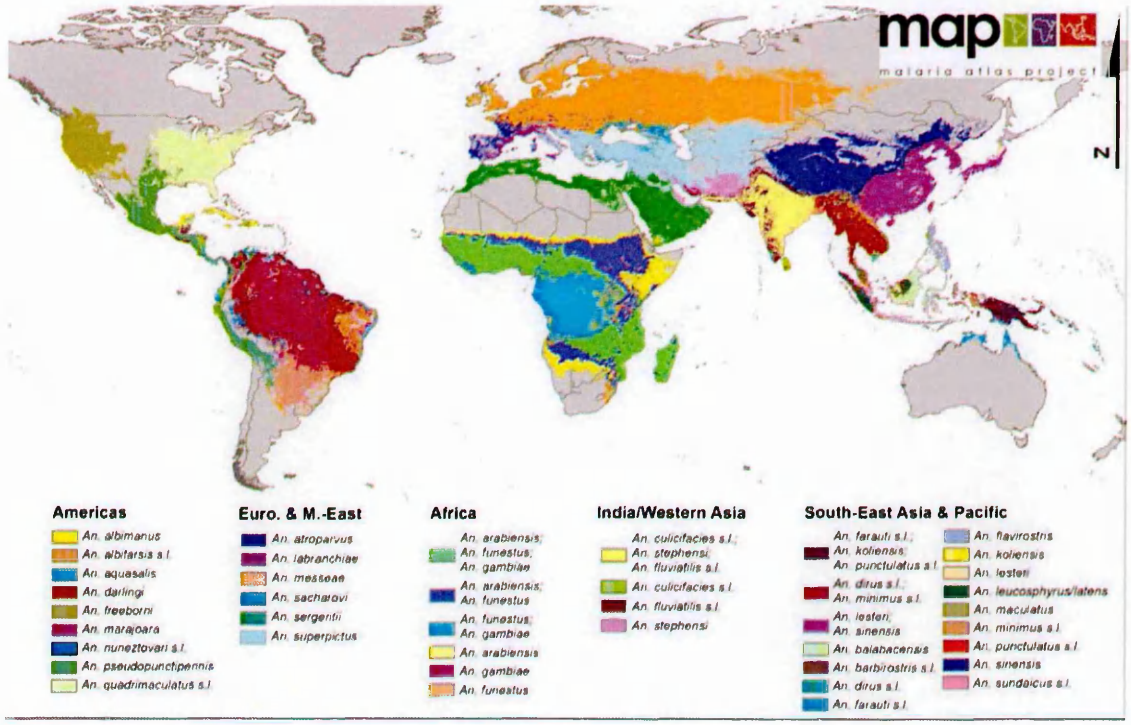
1.2.2 Vector

Mosquitoes of the Anopheles genus are responsible for malaria transmission. The genus consists of a total of 468 formally recognised species and more than 50 unnamed members of species complexes. Members of the genus differ in their geographical distribution and competence of transmission, with more dominant members being described as dominant vector species (DVS). Recently, information on their global distribution has been collated by the Malaria Atlas Project (Hay *et al.*, 2010; Sinka *et al.*, 2012), as shown in Figure 1.2. Although malaria has been eliminated from Europe and much of the Americas, dominant vector species still exist in these countries (Sinka

et al., 2012), therefore potential for transmission exists beyond malaria endemic countries.

Only female anopheles mosquitoes transmit malaria. This is because such mosquitoes fulfil two basic criteria for malaria transmission: firstly, they require a blood meal as a protein source for development of their eggs, and therefore feed on humans/animals and secondly; the parasite is able to infect them in the process of taking a blood meal and subsequently develop in the mosquitoes gut, allowing the parasite's transmission the next time the mosquito takes a blood meal. It is possible to quantify the vector's infectious rate using measures of entomological inoculation rate (EIR).

Figure 1. 2 Global distribution of dominant vector species of malaria. Adapted from (Sinka *et al.*, 2012)



A modelled map showing the areas where surveys have identified dominant vector species of malaria transmission globally.

1.2.3 Human host

Various factors influence malaria in the human host, including intensity of transmission, clinical presentation, malaria immunity, host genetics and their interactions and the age of the host population. These are considered briefly in the following sub-sections.

1.2.3.1 Intensity of malaria transmission

Common measures of malaria transmission intensity include:

(i) Spleen Rate

(ii) Parasite Rate (PR)

(iii) Entomological Inoculation Rate (EIR)

i) Spleen rate

Initial transmission metrics were based on spleen rates, which was essentially a measure of the prevalence of palpable spleens. Such endemic malaria classifications were first formalized by a World Health Organization (WHO) conference in 1950 (WHO, 1951) as follows:

Hypoendemic malaria: Spleen rate in children 2-10 years of age, 0-10%.

Mesoendemic malaria: Spleen rate in children 2-10 years, 11-50%.

Hyperendemic malaria: Speen rate in children 2-10 years, constantly over 50%, spleen rates in adults high.

Holoendemic malaria: Spleen rate in children 2-10 years, constantly over 75%, spleen rate in adults, low; it is in this type of endemicity that the strongest adult tolerance is found.

ii) Parasite rates

A more recent classification of malaria endemicity quantifies malaria risk based on the parasite rate (PR), which has been defined as the proportion of individuals infected during a survey.

The parasite rate has been used to define malaria risk classes (Hay *et al.*, 2008) as follows:

No risk areas: Parasite rate 0

Unstable risk *P. falciparum* annual parasite index (PfAPI) < 0.1 per 1,000 per annum;

Low risk *P. falciparum* Parasite Rate in children 2-10 years old ($PfPR_{2-10} \leq 5\%$);

Intermediate risk ($PfPR_{2-10}$ 5-40%);

High risk ($PfPR_{2-10} \geq 40\%$).

The 2-10 year category denotes children who bear the brunt of malaria infection.

iii) Entomological inoculation rate (EIR)

This is the most rigorous measure of transmission, and aims to estimate the average number of infectious bites received per person per year. EIR is calculated as a product of the vector biting rate and sporozoite rate (SR), which is the proportion of mosquitoes with sporozoites in their salivary glands (Sylla *et al.*, 2000). Due to the rigour involved, data on EIR are fewer, but remain the most accurate measure of transmission (Burkot & Graves, 1995). Despite their accuracy, EIR data points are sparse and therefore the generalizability of the measure faces challenges. As such, modelling approaches represent one way to increase the generalizability of EIR over geographical locations that have not been sampled (Amek *et al.*, 2012).

1.2.3.2 Clinical presentation of malaria in the host

Infection with malaria results in highly heterogeneous outcomes in the host, ranging from asymptomatic carriage of the parasite, mild disease or even life threatening or fatal disease. Factors affecting this progression include the intensity of transmission and host immune factors. In individuals with clinical symptoms, presentation can be broadly classified into two forms; uncomplicated malaria and severe malaria.

1.2.3.2.1 Uncomplicated malaria

Most episodes of malaria in endemic areas are not complicated by life-threatening features. Uncomplicated malaria episodes are commonly characterised by symptoms such as malaise, dizziness, fatigue, headaches, muscular and joint pains, diarrhoea and vomiting. Within 2-3 days the infected individual presents with characteristic paroxysms of fever that are the hallmark of malaria infection. Nevertheless, the clinical picture of malaria is difficult to distinguish from other childhood illnesses, and therefore a slide positive for malaria is usually required to confirm malaria in individuals presenting with such symptoms. However, many individuals in malaria endemic areas may have malaria parasites in their blood but remain asymptomatic, complicating the diagnosis of malaria even further. A threshold of parasites specific to a particular endemic population is often used to determine which individuals presenting with suspected malaria fevers are actually suffering from malaria (Smith *et al.*, 1994).

1.2.3.2.2 Severe malaria

A small subset of individuals develop life threatening complications during clinical episodes of malarial infection (Greenwood *et al.*, 1991). These are most common

during infections caused by *P. falciparum* (Organization, 2014; WHO, 2000), although there is growing evidence that life threatening malaria can also, on occasion result from infections with *P. vivax* (Baird, 2009). Such presentations of malaria are generally termed severe malaria, and include syndromes that include impaired consciousness, respiratory distress due to acidosis, severe anaemia, renal failure, hypoglycaemia and jaundice (Organization, 2014; WHO, 2000). The classical symptoms of severe malaria are not the same in all patients, as presentation varies with transmission settings, underlying conditions, and background immunity of the host (WHO, 2000). For example, in hyper-endemic areas, severe malarial anaemia is the most common complication of severe malaria encountered, whereas in areas of low endemicity, cerebral malaria is a more common presentation. It is worth noting that overlaps exist between the various conditions termed severe malaria (Marsh *et al.*, 1996), but these complications are conceptually understood as distinct syndromes such as impaired consciousness, respiratory distress, severe anaemia, hypoglycaemia and jaundice.

1.2.4 Host genetics and malaria

In many malaria holoendemic areas, young children experience several clinical attacks of malaria each year but only a small proportion of these attacks lead to life threatening illness (Greenwood *et al.*, 1991). As described above, malaria associated mortality is mostly due to severe malaria syndromes such as severe malarial anaemia, metabolic acidosis, seizures, prostration and coma. Human host genes have been shown to be important determinants of the progression from mild clinical attacks to severe and life threatening malaria (Greenwood *et al.*, 1991; Mackinnon *et al.*, 2005).

The most commonly distributed genotypes that appear to offer host protection from malaria are the ABO blood groups (Rowe *et al.*, 2009) and the haemoglobinopathies, such as haemoglobin S (HbS), haemoglobin C (HbC) and α^{+}thalassaemia. Although human populations vary widely in their genetic structure (Tishkoff *et al.*, 2009), malaria endemicity presents a common selection pressure in selection for particular protective variants of genes, and can predict the gene frequencies in populations as shown by global evidence based maps of HbC (Piel *et al.*, 2013), HbS (Piel *et al.*, 2010), glucose-6-phosphate dehydrogenase (G6PD), deficiency (Howes *et al.*, 2012) and Duffy antigen negativity (Howes *et al.*, 2011).

A contemporary meta-analysis of the effects of haemoglobinopathies such as HbS, α^{+}-thalassaemia and HbC on malaria has recently been conducted by Taylor and colleagues (Taylor *et al.*, 2012). Common genes conferring malaria resistance are considered below.

1.2.4.1 Classical malaria protective genes

1.2.4.1.1 Sickle cell trait

The malaria protective effect of sickle cell trait (HbAS) was first suspected over 60 years ago (Allison, 1954; Beet, 1946). It is now clear that HbAS confers protection against all forms of clinical malaria including severe malarial anaemia (SMA), cerebral malaria (CM) and malaria related convulsions (May *et al.*, 2007; Williams, Mwangi, Wambua, Alexander, *et al.*, 2005). Furthermore, the fewer cases of malaria in HbAS individuals are associated with lower parasite densities than in normal (HbAA) individuals (Williams, Mwangi, Wambua, Alexander, *et al.*, 2005). The mechanism of protection afforded by HbAS has been of interest and while some

reports attribute it to innate mechanisms (Friedman, 1978; Friedman *et al.*, 1979; Pasvol *et al.*, 1978; Roth *et al.*, 1978), the adaptive immune mechanism has been implicated in some *in vitro* studies which show higher lympho-proliferative responses to *P. falciparum* soluble antigens in children with HbAS compared with children of normal genotype, HbAA (Abu-Zeid *et al.*, 1992; Bayoumi, 1987). Regarding antibody-mediated protection, studies demonstrated higher titres of antibodies in HbAS compared with normal individuals (Marsh *et al.*, 1989). Another observational study found evidence of a possible interaction between trans-placentally acquired antibodies and HbAS in the delay of onset of clinical malaria in new-borns. This protection was only present when HbAS and the antibodies were present together, and was abolished when they were present separately, indicating an interactive effect of HbAS in antibody-mediated immune protection (Achidi *et al.*, 1996). There is a scarcity of population genetic epidemiology studies to assess differences in immune protection between normal and HbAS individuals. One such study found differences in lymphocyte proliferative responses to *P. falciparum* antigens between normal and HbAS children, which varied with age (Le Hesran *et al.*, 1999). Similarly, observational age-stratified longitudinal cohort studies conducted in Kilifi, Kenya have shown that the protection afforded by HbAS against malaria is not the same in different age categories. Incidence rate ratios of HbAS compared to HbAA show protection from malaria increasing from only 20% in children <2 years to a peak of 56% at 8-10 years of age, strongly arguing for an immune component in HbAS mediated protection (Williams, Mwangi, Roberts, *et al.*, 2005). As such, the current study aimed to document the age-specific susceptibility to malaria in HbAS children to investigate further the evidence of adaptive immunity in malaria protection conferred by HbAS.

1.2.4.1.2 α-thalassaemia

α-thalassaemia results in a deficiency in the α-chains of the haemoglobin molecule due to deletions or mutations in one or more genes encoding these chains. Previous studies have shown that α-thalassaemia, results in as much as 40% protection from severe malaria anaemia (Wambua *et al.*, 2006). As in the case with HbAS, *in vitro* studies suggest possible mechanisms of α-thalassaemia protection. Yuthavong (Yuthavong *et al.*, 1988) demonstrated increased susceptibility of infected α-thalassaemic red blood cells (RBCs) to phagocytosis while others have shown altered surface antigen expression (Luzzi *et al.*, 1991; Udomsangpetch *et al.*, 1993), a phenomenon that may be relevant to immunogenicity and pathogenicity. Other reports showed a common mechanism of enhanced phagocytosis in ring-parasitized mutant RBCs with HbAS, G6PD deficiency and β-thalassaemia, however, this was not the case in α-thalassaemic RBCs suggesting a different mechanism of protection from malaria (Ayi *et al.*, 2004; Cappadoro *et al.*, 1998).

A genetic epidemiology study in the Pacific found higher susceptibility to mild malaria in children with α-thalassaemia relative to normal children, suggesting that protection conferred against severe malaria by α-thalassaemia might be due to priming of the immune system by the mild malarial infections (Williams *et al.*, 1996). In contrast, another study on the Kenyan coast did not find evidence of increased risk of mild malaria in α-thalassaemic children but instead demonstrated a trend towards a reduced incidence of uncomplicated malaria (Wambua *et al.*, 2006). In this study, α-thalassaemia protection against severe malaria was specific to SMA. A similar observation was made in a study conducted in Ghanaian children (May *et al.*, 2007). Finally, in another population-based study, humoral responses in children with α-

thalassaemia were no different from those in normal individuals (Allen *et al.*, 1993). As such, there is a need for further population-based genetic studies to delineate the precise association between α-thalassaemia and susceptibility to severe malaria.

1.2.4.1.3 Complement receptor 1 (CR1) polymorphisms

The Swain Langley (Sl) and McCoy (McC) genotypes of red cell complement receptor 1 (CR1) have been associated with protection from CM (Thathy *et al.*, 2005). In addition, one hypothesis implicated in the protection of α-thalassaemia against severe malaria has been demonstrated to be through reduced expression of the CR1 in α-thalassaemic RBCs. This may partly be due to the *in vitro* rosetting property of the CR1 protein (Rowe *et al.*, 1997). Rosetting has been associated with severe malaria in a number of studies in Africa (Stoute, 2005) and therefore population genetic studies of polymorphisms in the CR1 gene may provide an opportunity for exploring the role of CR1 in the pathophysiology of severe malaria.

1.2.4.1.4 G6PD deficiency

G6PD is a key enzyme in the pentose phosphate pathway that protects RBCs from oxidant damage. G6PD deficiency is an enzymopathy that has been known to protect against malaria for about 50 years (Allison & Clyde, 1961). It is an X chromosome-linked condition, meaning males, with only one X chromosome, are hemizygous whereas females, who have two X chromosomes, can either be homozygous or heterozygous for the deficiency. The non-deficient genotype is referred to as G6PD B. Different allelic variants may cause deficiency, but in Kenya, the mutant responsible for deficiency is the 202 C>T mutation, referred to as G6PD A- (Carter *et al.*, 2011; Shah *et al.*, 2014). This A- mutation (202C>T) is associated with approximately 40%

of the normal G6PD B enzyme activity, and only occurs in the background of the 376 (T>C) mutation. Individuals with only the 376 T>C mutation have G6PD A, which is associated with 85% enzyme activity of G6PD B. The role of G6PD deficiency in protection from malaria remains controversial since many reports attribute different levels of protection (Allison & Clyde, 1961; Bienzle *et al.*, 1972; T. Butler, 1973; Gilles *et al.*, 1967; Guindo *et al.*, 2007; Ruwende *et al.*, 1995). However, since these studies were heterogeneous, they are difficult to compare especially in the light of the different diagnostic methods used and the different outcomes assessed for protection. Recently, the debate on mechanisms of protection has been informed by two contrasting studies suggesting different mechanisms. In one study, both hemizygotes and heterozygotes were protected from severe malaria (Ruwende *et al.*, 1995) whereas in another investigation, only the hemizygous males were protected but not heterozygous females (Guindo *et al.*, 2007). It is evident that more population-based epidemiological studies are required to inform the debate on mechanisms of G6PD deficiency protection.

1.2.4.2 Recently identified malaria-protective polymorphisms

Classical malaria resistance genes like HbAS have been shown to account for only a small fraction of genetic resistance to severe malaria, suggesting that many other genes act out to protect individuals from malaria (Mackinnon *et al.*, 2005). These have been the subject of association studies to identify novel genes and confirm replication of protection across populations (MALGEN, 2008). Recently, genome wide association studies have succeeded in identifying 2 new SNPs with reproducible protection across populations, the rs4951074 at chromosome 1q32.1, ATP2B4 locus and the rs2334880 at chromosome 16q22.2 linked to the MARVELD3 locus (Timmann *et al.*, 2012). The statistical association between these genes and differential malaria outcomes is

augmented by the functional role of these loci. ATP2B4 locus encodes the major calcium pump in erythrocytes, which are the main pathogenic host cells, whereas MARVELD3 is a protein expressed on endothelial cells and as such sequestration of parasitized erythrocytes may be aided by variants of MARVELD3. Many of the genetic associations have recently been confirmed in large multicentre studies (MALGEN, 2014).

1.2.4.3 Human genetic studies may lead to malaria interventions

Based on an understanding of the malaria lifecycle (see Figure 1.2), interventions against malaria commonly target the parasite, vector or the human host risk factors. Parasite and vector targeting is quite recent, meaning that human host genetic factors have been the 'natural interventions' or determinants of malaria outcomes for millennia. For example, the use of quinine, which kills the parasite in the blood stages, has been documented in the last 400 years (circa 1600) in the Andes ranges by Jesuit missionaries (A. R. Butler *et al.*, 2010), though ancient populations probably used it for millennia (Waterfall CM, 2001). Vector based interventions that are aggressive and of a large scale have targeted the mosquito since the discovery of mosquitoes as the vectors for malaria transmission by Ronald Ross in 1898 (R. Ross, 1897). Human genetic factors protecting from malaria are likely to have played a vital role in distinguishing between survival and death from malaria in the millennia before the introduction of chemotherapy and vector control.

Studies of human genetic polymorphisms represent a natural experiment of which genetic make-ups are better suited to survive the malaria challenge. Such studies can result in interventions. For example the discovery that Duffy antigen negativity makes

red blood cells refractive to *P. vivax* malaria (Miller *et al.*, 1976) infection has led to the identification of a parasite protein, the *P. vivax* Duffy binding protein (PvDBP) as a major target for vaccine development (Chitnis & Sharma, 2008). This discovery has also led to better descriptions of populations at risk of infection, for example *P. vivax* is delimited from most of Africa because the majority of the African population is Duffy antigen negative (Guerra *et al.*, 2010). Recent evidence suggests that *P. vivax* originated in Africa (Liu *et al.*, 2014) and led to the selection of the Duffy negative allele to fixation (Hamblin *et al.*, 2002). Similarly, the discovery of the association between Human Leukocyte Antigen (HLA) type and the risk of severe *P. falciparum* malaria (Hill *et al.*, 1991) has provided a rationale for the development of vaccines that target the pre-erythrocytic stages of parasite development by inducing T-cell responses.

Another example of how genetic polymorphisms interact with interventions is offered by G6PD deficiency; individuals harbouring deficiency associated mutations experience haemolysis when treated using primaquine (Howes *et al.*, 2012). The numbers of individuals carrying such mutations are an important measure when considering this treatment option, especially in the current age of malaria elimination since it is the only available drug in the pharmacopeia that can both prevent transmission by attacking gametocytes and preclude relapse by killing hypnozoites that may otherwise relapse (Eziefula *et al.*, 2012). Careful characterization of G6PD deficient populations in such campaigns may lead to better deployment and delineation of where such campaigns are likely to be most effective.

The classical human host genes associated with malaria resistance affect the RBC where the malaria parasite mainly resides in its life cycle in the human host (Williams,

2006). So far, mechanistic studies have shown that some of these conditions are associated with enhanced production of antibodies directed towards the infected red cell surface that enhance their destruction through phagocytosis. This raises the possibility of using agents that mimic the antigenic presentation caused by these conditions as new therapeutic agents (Kennedy, 2001).

1.2.4.4 Age- and genotype-specific malaria incidence

1.2.4.4.1 Relationship between genetic variants and age patterns of disease

Severe malarial disease and mortality predominate among the youngest children (WHO, 2013) and tapers off sharply in the older individuals in endemic populations. In order for genetic variants that confer protection to achieve high population frequencies, they would need to protect from death at the youngest ages that bear the highest mortality burden and be fit enough for passage to future generations. The classical gene associated with malaria protection, HbAS, meets these criteria, as epidemiological evidence shows it protects from the most severe malaria symptom among those under 5 years old (Williams, Mwangi, Wambua, Alexander, *et al.*, 2005) and allows survival from life threatening malaria. Survivors with the carrier status are able to pass it on to their offspring, with the continual malaria selective pressure in succeeding generations leading to more carriers of the gene surviving to propagate the gene. Thus, the gene achieves high population frequencies as a result of the protective effect that it confers against malaria (Piel *et al.*, 2010).

However, Mackinnon and colleagues (Mackinnon *et al.*, 2005) found that the HbAS could only account for 2% of the heritability of protection from malaria, out of a total genetic heritability of approximately 25%. Given that other canonical examples of

malaria protective genes may not account for more than 2% of genetic heritability, one would expect that approximately 90% of genetic heritability of malaria protection remains to be explored. Mackinnon and colleagues concluded that most likely, many monogenic causes each contribute a small fraction to overall susceptibility. More recently, Alcais and colleagues (Alcais *et al.*, 2010) have proposed that single gene variations underlie the severity of paediatric infectious diseases and that this effect is reflected in the age specific incidence patterns of various variants of protective genes.

According to Alcais and colleagues (Alcais *et al.*, 2010), infectious diseases show characteristic age profiles of morbidity and mortality, with each pathogen showing a characteristic age pattern of morbidity/mortality. Examples include the U shaped pattern of higher disease in the youngest and oldest individuals with an intervening dip in adolescents and the middle aged as seen in invasive pneumococcal disease and herpes simplex encephalitis; and the L shaped pattern of higher disease in only the youngest children, as seen in diseases such as malaria, measles, tetanus and in diphtheria and poliomyelitis before the widespread use of vaccination. These patterns are sensitive to changes in the pathogen or the host. For the pathogen, peak shifts in the burden of disease may represents a change in the pathogen to more virulence or higher transmission, examples being pandemic influenza and Ebola (Alcais *et al.*, 2010).

Alternatively, a shift in the age burden could reflect immunological changes in the host. In the case of malarial disease, the characteristic L shape of the age specific incidence curve occurs because the youngest children bear the highest burden of clinical malaria as they lack adaptive immunity. As they age and acquire immunity,

clinical disease sharply tapers off to a more uniform risk after puberty (Baird, 1995). This age incidence curve is characteristic of endemic populations. However, in non-endemic populations, the majority of the population is immunologically naive to the pathogen and the incidence pattern may be more similar across the population.

In *P. falciparum,* changes in transmission intensity shift the peak prevalence of infection to younger age groups as transmission intensity increases (Molineaux L, 1980). Acquisition of immunity in the host is associated with the peak shift phenomenon (Woolhouse, 1998). Conversely, the age incidence curves normally shift to older age groups in response to interventions, because the interventions reduce exposure, thereby delaying the development of acquired immunity. It has been observed that declines in malaria transmission result in shift of malaria morbidity to older children, especially in the current push towards malaria elimination (Ceesay *et al.*, 2008; O'Meara *et al.*, 2008; Schellenberg *et al.*, 2004). Importantly, the peak shift phenomenon is not limited to malaria, but is also seen in other infections where immunity is acquired (Woolhouse, 1998).

Knowledge of the patterns of age-malaria incidence has been used to target interventions to the worst affected, children less than 5 years in the case of malaria (WHO, 2013). The age-specific incidence of malaria is therefore an important tool in monitoring the shifting burden of malaria as transmission intensity declines, and allows an appropriate response to clear reservoirs of malaria as transmission declines (WHO, 2013).

1.2.4.4. 2 Age-specific incidence of malaria by genotype

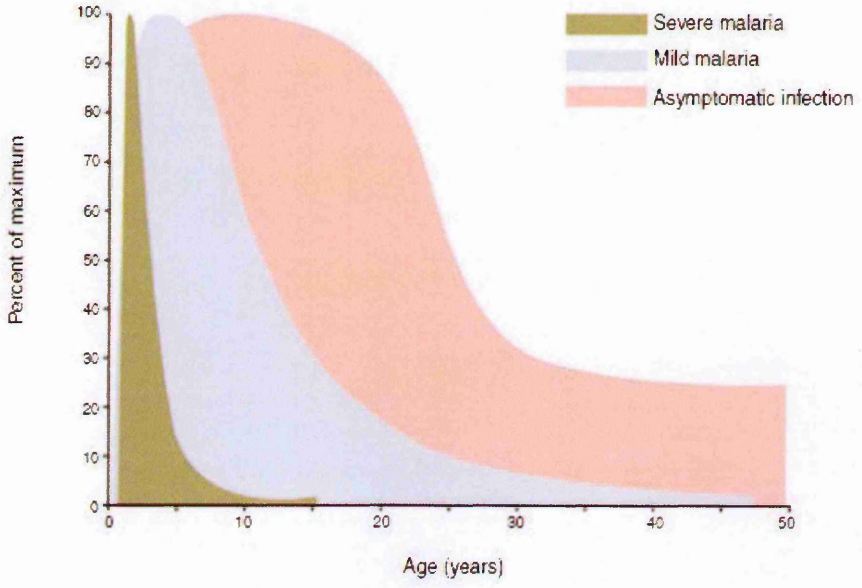
From the foregoing, it is clear that the age-incidence curve is a proxy for the rate of acquisition of naturally acquired immunity. It is conceivable that genetic factors protecting from malaria can have the same effect as the protective interventions and may shift the age curves to older age groups. This has previously been investigated in an age cohort by studying the effect of protection afforded by the sickle cell trait across a variety of various ages (Williams, Mwangi, Roberts, *et al.*, 2005). Williams and colleagues argued that if immune mechanisms are involved in the protection afforded by the sickle cell trait, the degree of protection should increase up to the age when children would normally become functionally immune to malaria, and that any immunological advantage beyond this time would not result in any increased level of protection. It was shown that the level of protection increased approximately threefold, from only 20% in the first 2 years of life to about 56% in those who were 10 years of age and thereafter a reduction to 30% in those older than 10 years. This was therefore consistent with an immune basis of protection.

In another study, Gong and colleagues (Gong *et al.*, 2012) followed a cohort of children 1-10 years of age and compared their establishment of parasitaemia, symptomatic malaria and parasite density. Sickle trait carriers were protected against the establishment of parasitaemia and development of symptomatic malaria at older ages but not at younger ages. However, the younger children were protected against high parasite density compared with the older children in the cohort. Effects only seen at older ages suggest acquired mechanisms of protection, compared to effects that are seen at younger ages, which suggest innate mechanisms of protection. Thus, it appears that the protection afforded by sickle cell trait has both innate and acquired components.

Age incidence patterns can also give insights into possible mechanisms of protection associated with immunity. For example, in a study of age specific incidence of malaria in the pacific (Williams *et al.*, 1996), young α-thalassaemia homozygous (-α/-α) children (< 4 years old) had more clinical malaria than their normal (αα/αα) counterparts. The effect was not seen in those who were 5-9 years. This was paradoxical, as α^{+}thalassaemia is known to protect against malaria. The authors concluded that α^{+}thalassaemia may beneficially predispose to malaria in early life to avoid later severe disease, what they termed a 'natural' vaccine i.e. a genetic condition that predisposes to disease in early life in order to avoid the severe symptoms of the disease in later life. The authors argue that since diseases such as rubella and poliomyelitis cause least morbidity when contracted early in childhood, this is evidence of priming the immune system through exposure to less severe episodes to avoid the life threatening illnesses in later life. Interestingly, when I considered the poliomyelitis age incidence curve reported by Alcais (Alcais *et al.*, 2010), I observed that it is L shaped, similar to that for malaria. This led me to hypothesize that 'natural vaccination' may be conferred by different genes and may be replicated in different populations or transmission settings. This study investigates this by observing age- and genotype-specific malaria incidence in Lwak, Asembo area of Western Kenya. A characteristic age incidence pattern of malarial diseases is illustrated in Figure 1.3 below:

Figure 1. 3 Characteristic age incidence patterns of severe, mild and asymptomatic malaria.

Adapted from (Langhorne *et al.*, 2008).



The figure shows how indices of malaria change over time in an endemic area. Green represents severe malaria, most prevalent in the children the least exposed members (under 5 years) whereas purple shows the proportion of mild malaria in older children, adolescents and young adults, and while orange shows asymptomatic infection most prevalent in the most exposed members of the population, the oldest.

Through this study, I will seek evidence for age and genotype-specific changes in peaks of infection, consistent with different rates of acquired immunity in different ages and genotypes respectively. If stratification by the protective variants of the genes shows a difference in peak shift towards older children, this provides evidence of the role of immunity in protection, whereas if stratification by the protective genes does not show a peak shift, immunity may play no role in the protection afforded by these genes.

1.2.4.5 Role of epistasis in malaria incidence studies

Epistasis is simply interaction between loci that predicts the phenotype. Carlborg and colleagues (Carlborg & Haley, 2004) have reported that it is more common than previously thought, and should be considered in genetic studies. For example, it has recently been observed that the separate malaria-protective effects of HbAS and α-thalassaemia are each abrogated when these conditions are inherited together (Williams, Mwangi, Wambua, Peto, *et al.*, 2005). Together with the previous observations regarding HbAS and α-thalassaemia above, this finding suggests contrasting mechanisms of malaria protection for HbAS and α-thalassaemia. In co-inheritance between HbAS and α-thalassaemia, there is reduced synthesis of HbS due to a lower availability of α chains to combine with β^s chains (Mouele *et al.*, 2000). Since the HbS component in the sickle cell trait individuals protects from malaria, when it is reduced due to co-inheritance with α-thalassaemia, it follows that the protective effect will be reduced. Similarly, the observation that there is epistasis between α-thalassaemia and haptoglobin (Atkinson *et al.*, 2014) led to inferences on the mechanism of the interaction between α-thalassaemia and the haptoglobin 2-2 molecule. These include the following:

(i) Hp2-2 is less able to quench the unmatched globins in α-thalassaemia leading to increased oxidative stress and therefore reducing the protection afforded by α-thalassaemia to malaria.

(ii) Both α-thalassaemia and haptoglobin lead to increased methemoglobin. Since methemoglobin, is an endothelial cell activator, the effect of excess methemoglobin may reduce protection from severe effects of malaria.

(iii) Hp2-2 increases pro-inflammatory response thus discounting any protection by α-thalassaemia.

Studies of epistasis can thus lead to possible new avenues of inquiry into the mechanisms of malaria protection afforded by malaria protective genes.

The current study investigates malaria epidemiology and the effect of malaria candidate genes and gene interactions on malaria incidence in Asembo, western Kenya.

1.2.4.6 Entomological profile and malaria transmission in Asembo

Historical trends of EIR in Asembo are associated with changes in vector populations, proportions and their species-specific EIRs. In the late 1990s, following a large scale randomized trial which distributed insecticide treated nets (ITNs) to residents, populations of vector species diminished, and sporozoite rates declined by approximately 90% (Gimnig *et al.*, 2003). Whereas *An. gambiae* and *An. funestus* were the dominant vectors before the ITN intervention (Gimnig *et al.*, 2003), in the mid to late 2000s, *An. arabiensis* became the predominant vector while low proportions of *An. gambiae* and *An. funestus* were reported (Bayoh *et al.*, 2010). However, the *An. funestus* population has resurged recently (McCann *et al.*, 2014).

Measures of EIR have changed over the same period. Compared to data collected before the large scale distribution of ITNs, daily EIRs have shown a consistent reduction for An. gambiae *sensu latu* (indoor EIR = 0.432 in 1985–1988, 0.458 in 1989–1990, 0.023 in 2011). An. *a*rabiensis, a member of the *An. gambiae sensu latu* complex (the other being *An. gambiae ss*) has shown an even more marked reduction (indoor EIR = 0.532 in 1989–1990, 0.039 in 2009, 0.006 in 2011). However, An. funestus has shown an overall increase in EIR (indoor EIR = 0.029 in 1985–1988, 0.147 in 1989–1990, 0.010 in 2009 and 0.103 in 2011) (Adazu *et al.*, 2005; Bayoh *et*

al., 2014) Malaria transmission occurs year round but EIR peaks in May during the wet season each year (Amek *et al.*, 2012).

1.2.4.7 Parasite prevalence reported in Asembo

P. falciparum prevalence in Asembo reduced from a high of >80% in the 1990s (Bloland *et al.*, 1999) to <30% in 2008, and are currently at about 40% since 2009 (Hamel *et al.*, 2011). The overall reduction from the 1990s prevalence followed the introduction of insecticide treated nets (ITN) scale up and a policy shift from SP to artemisinin based combination therapy (Amin *et al.*, 2007). Parasite prevalence within the KEMRI CDC HDSS is not uniform, with recent reports showing that the 5-14 year age category has the highest parasite prevalence of 58% compared to the <5 years and the ≥15 years (KEMRI CDC unpublished data, referenced in (Desai *et al.*, 2014)). According to a model-based map (Noor *et al.*, 2009), *P. falciparum* malaria transmission in Asembo is among the highest in the country, with parasite prevalence recorded at ≥40% in children aged 2-10 years. However, such surveys use rapid diagnostic tests (RDTs) for the diagnosis of malaria, which has the potential to overestimate the malaria positive rate, as reported by Aydin-Schmidt and colleagues (Aydin-Schmidt *et al.*, 2013). This is because RDTs target the parasite antigen histidine rich protein 2 (HRP2), which persists in the blood even after successful treatment. RDTs are often favoured for malaria surveys because they are logistically easier to deploy. However, recent data has shown that polymerase chain reaction (PCR) based methods have the ability to detect parasites at much lower levels than those detected by microscopy or RDT (Kipanga *et al.*, 2014), while avoiding the false positive rates associated with persisting parasite antigens in some RDT methods.

1.2.4.8 Aims and objectives of the human genetic study in Asembo, western Kenya

Through this thesis, I aim to investigate the epidemiology of malaria in a birth and age cohort of children stratified by genetic polymorphisms including the HbAS, G6PD deficiency, CR1 and α-thalassaemia. These polymorphisms are expected to exist at high frequencies in the study area since previous studies conducted locally found frequencies of 17% for HbAS (Aidoo *et al.*, 2002), 44% for α^{+}thalassaemia heterozygotes (Kifude & Waitumbi, 2007) and 10% for G6PD deficiency (A-) (Hunja, 2006). The CR1 genotypes associated with protection, Swain Langley 2/2 and McCoy a/b were each found at 45% frequency (Thathy *et al.*, 2005). ABO blood groups will also be typed as possible confounders in malaria association analyses. In addition, the effect of co-inheritance of these genes, which is rarely analysed in such genetic epidemiology studies, will be investigated. Other genes involved in malaria protection will also be investigated by typing for Single Nucleotide Polymorphisms (SNPs) in 34 malaria candidate genes (see Appendix III).

The specific objectives of this thesis are:

i) To describe the epidemiology of malaria in comparison with common infectious childhood illnesses in Asembo, western Kenya.

ii) To investigate the effect of haemoglobinopathies (HbAS, G6PD deficiency, α-thalassaemia) and common red blood cell polymorphisms (ABO, Sl and McC blood groups) on malaria incidence.

iii) To investigate the effect of polymorphisms in 34 additional malaria candidate genes on malaria incidence.

iv) To show the effect of interactions between genes on malaria incidence.

The proceeding chapters describe the methods used and gives results from investigations of each of these objectives. These chapters are therefore organized as follows:

Chapter 2 outlines the materials and methods employed for the above investigations.

Chapter 3 describes the epidemiology of malaria in Asembo, western Kenya, in the context of other infectious disease syndromes detected in the area.

Chapter 4 investigates the effect of red blood cell polymorphisms on malaria incidence.

Chapter 5 investigates the effect of polymorphisms in 34 additional malaria candidate genes on malaria.

Chapter 6 investigates possible interactions between genes and their effect on malaria incidence.

Chapter 7 presents a discussion and conclusion of findings from the above investigations and suggests future directions for further work.

CHAPTER 2: Materials and Methods

2.1 Study site

The study was conducted in Asembo, part of the Rarieda Sub-County of Siaya County. Asembo borders Lake Victoria to the West, and covers an area of approximately 225 km^2. The population is culturally and ethnically homogenous, more than 95% being of the Luo ethnic group and living in dispersed rural settlements (Adazu *et al.*, 2005). The main economic activities are agriculture, fishing and small-scale businesses. The Sub-County is one of the poorest in Kenya, 70% of the population living on <1 USD / day (Bigogo *et al.*, 2010).

The Asembo area is one of the areas covered by a Health and Demographic Surveillance System (HDSS) that has been run by the Kenya Medical Research Institute (KEMRI) and Centres for Disease Control and Prevention (CDC) since 2001 (Figure 2.1). Within the Asembo HDSS area, the International Emerging Infections Program (IEIP) established an infectious disease surveillance system in 2005. The study participants for the current study were drawn from this surveillance area. The IEIP surveillance system is described in further detail in section 2.3 below.

2.2 Study population

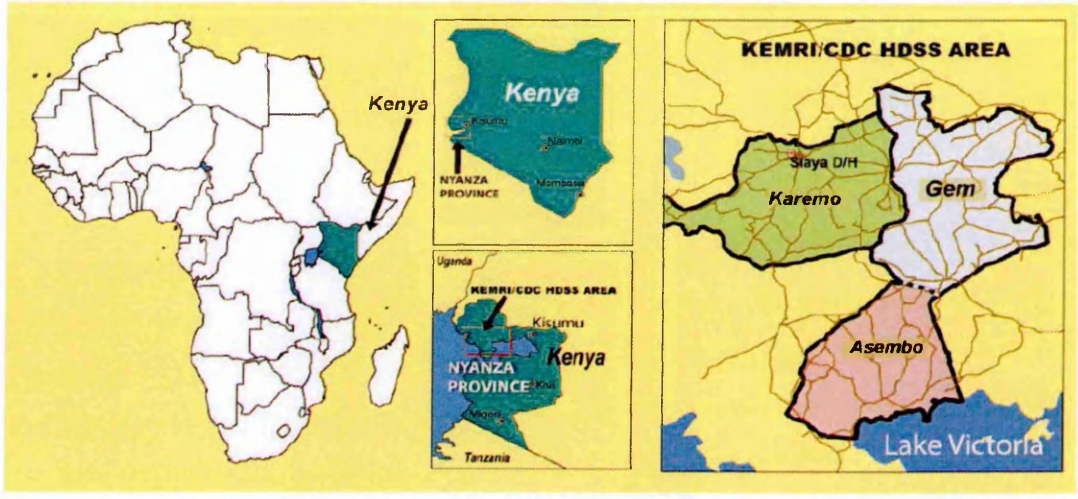
The study population is predominantly of the Luo ethnic group. The Luo are classified linguistically as Nilotes, and further sub-classified as River-Lake Nilotes because of their historically intimate association with rivers and with Lake Victoria. Originating from Sudan, a historically malaria endemic area, their migratory path along the river Nile and subsequent settlement on the shores of Lake Victoria in East Africa (Ogot, 1967) implies exposure to malaria in the lowlands for several millennia. In the absence

of aggressive and large-scale interventions against malaria except in the last century, the only survival advantage against malaria in this community is likely to be related to their behaviour or to human genetic factors that confer protection from malaria. Enrichment for frequencies of such factors would thus be expected among the Luo and as such, they are an excellent group to study for malaria-protective genetic factors.

2.3 International Emerging Infections Program (IEIP) on surveillance.

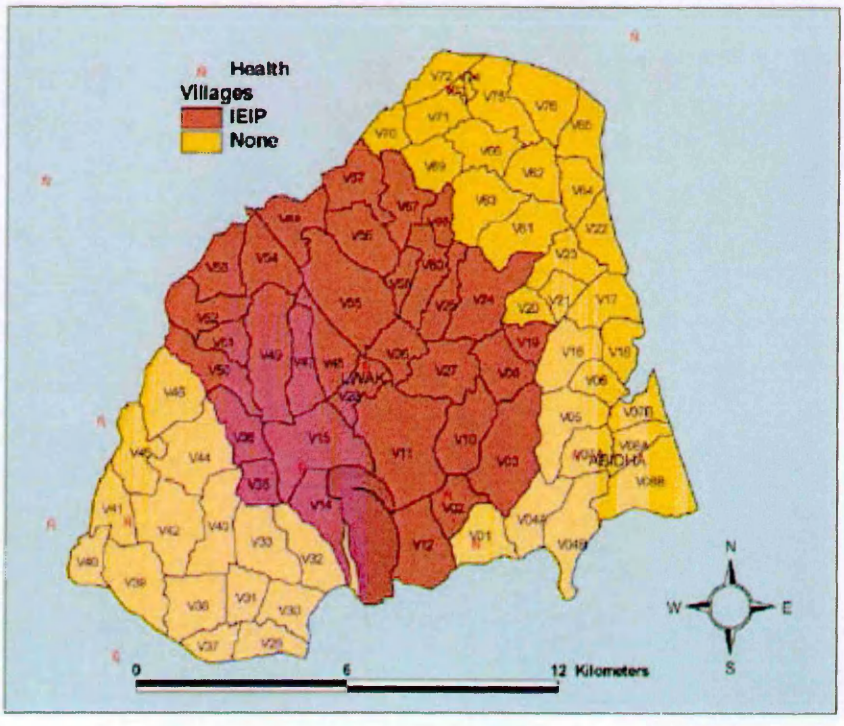
The IEIP has a programme of surveillance called PBIDS (Population Based Infectious Disease Surveillance), embedded within the KEMRI/CDC Health and Demographic Surveillance System (HDSS) (see figure 2.1). It is in the Asembo area of the HDSS, and covers an area of approximately 100km^2 in villages within a 5km radius of Lwak Mission hospital (Figure 2.2).

Figure 2. 1 The Asembo study areas in relation to Nyanza Province and Kenya.



Map adapted from Amek and colleagues (Amek *et al.*, 2012).

Figure 2. 2 Map of IEIP area of Asembo, part of the KEMRI/CDC HDSS



Adapted from (Bigogo *et al.*, 2010). The brown area is the IEIP surveillance area, whereas the yellow area is the HDSS area that is outside the IEIP.

PBIDS is mainly focussed on three infectious disease syndromes, namely acute respiratory infections, diarrhoeal diseases, and febrile illnesses. The catchment population included in PBIDS includes 33 villages in which at least 50% of the population live within a 5 km radius of Lwak Mission Hospital, the designated referral health facility. For a participant to be eligible for inclusion in the PBIDS, they had to be registered as residents of the HDSS run by KEMRI and CDC. HDSS residency is defined as being resident in the area for 4 continuous months i.e. not having moved from the study area for the previous 4 months.

Such surveillance may give useful data on disease syndromes in the area. An example is that home surveillance records between 2006 and 2008 indicated that health seeking for symptoms such as fever, diarrhoea and respiratory illness among children under-5 years ranged from 14-20% (Bigogo *et al.*, 2010).

2.3.1 Syndromic health and morbidity surveillance

As part of the main PBIDS platform, study participants are visited in their homes every two weeks and are interviewed to determine whether they are suffering from any one of a range of syndromes including respiratory illnesses, diarrhoea and fever. Those presenting with such symptoms are referred to Lwak Mission Hospital, where treatment is offered free of charge. The syndrome definitions used in the household visits are summarised in Table 2.1 below.

Table 2. 1 Definitions of syndromes used in the home morbidity surveillance conducted every two weeks.

Syndrome	Definition
Febrile illness	Illness in which subjective or documented fever was reported.
Diarrhoeal illness	An illness in which at least three looser than normal stools were reported during a 24-h period.
Acute Respiratory Illness (ARI)	Presence of a cough or difficulty in breathing.
Acute Lower Respiratory Illness (ALRI) in children:	Presence of cough or difficulty in breathing and either elevated respiratory rate for age or chest in-drawing on exam by trained field workers.
Acute Lower Respiratory Illness (ALRI) in adults	Presence of subjective or documented fever with cough, difficulty in breathing, or chest pain.

2.3.2 Passive clinic surveillance

Participants from within the study area are seen during inter-current episodes of illness at the designated surveillance hospital, the Lwak Mission Hospital. Details of the clinical procedures performed at the hospital have been published previously (Feikin *et al.*, 2011). Briefly, vital signs and symptoms are recorded by trained health facility recorders or nurses. These include axillary temperatures, respiratory rates (measured for 1 minute using an audible countdown timer) and oxygen saturation (measured using a fingertip pulse oximeter: NONIN Medical, Minnesota).

Blood smears are prepared for malaria diagnosis on participants presenting with a history of fever or who have a documented axillary temperature of >=38^{0}C. Qualitative

slide reading is done at the hospital to facilitate malaria case management. As such slides are declared either positive or negative and no quantitative malaria slide reading is done on routine visits. For particular ALRI patients who meet clinical indications, combined nasopharyngeal and oropharyngeal swabs are taken for PCR. Stool and blood samples are also collected for culture from patients meeting specific clinical indications. Structured questionnaires are used by clinicians to record details of sickness episodes including demographic data, case history and a summary of any drugs that may have been taken prior to the hospital visit. Between 2006 and 2013, scan-able paper forms (TeleForm® software, Cardiff ™, Vista, CA) were used at Lwak to record clinical details (See Appendix VII). However, a patient care system was adopted in 2013 in which data are entered directly into a custom-designed computerized system with internal quality control checks.

2.4 Human genetic factors and childhood disease study

Through this study I aimed to investigate the following questions:

- Does the incidence of malaria and other childhood infections differ across children of different genotypes?
- Do interactions between different genotypes affect the incidence rate of malaria and other childhood diseases?

The study design that was best suited to capture rates of malaria was a longitudinal cohort study, since the cohort may then be stratified into genotypes to give the genotype-specific rates of malarial and common childhood illness. This was satisfied by having a birth cohort which would allow observation of episodes from the youngest age, whereas an age cohort involving individuals of a wider age range allowing stratification over wider age bands. The final cohort recruited therefore consisted of

individuals followed up since birth, referred to as the birth cohort and older individuals who had follow-up data available since the year 2008, referred to as an age cohort.

My study was therefore conducted in a sequence of four steps as follows:

(i) Recruitment of study participants.

(ii) Genotyping for different malaria candidate genes.

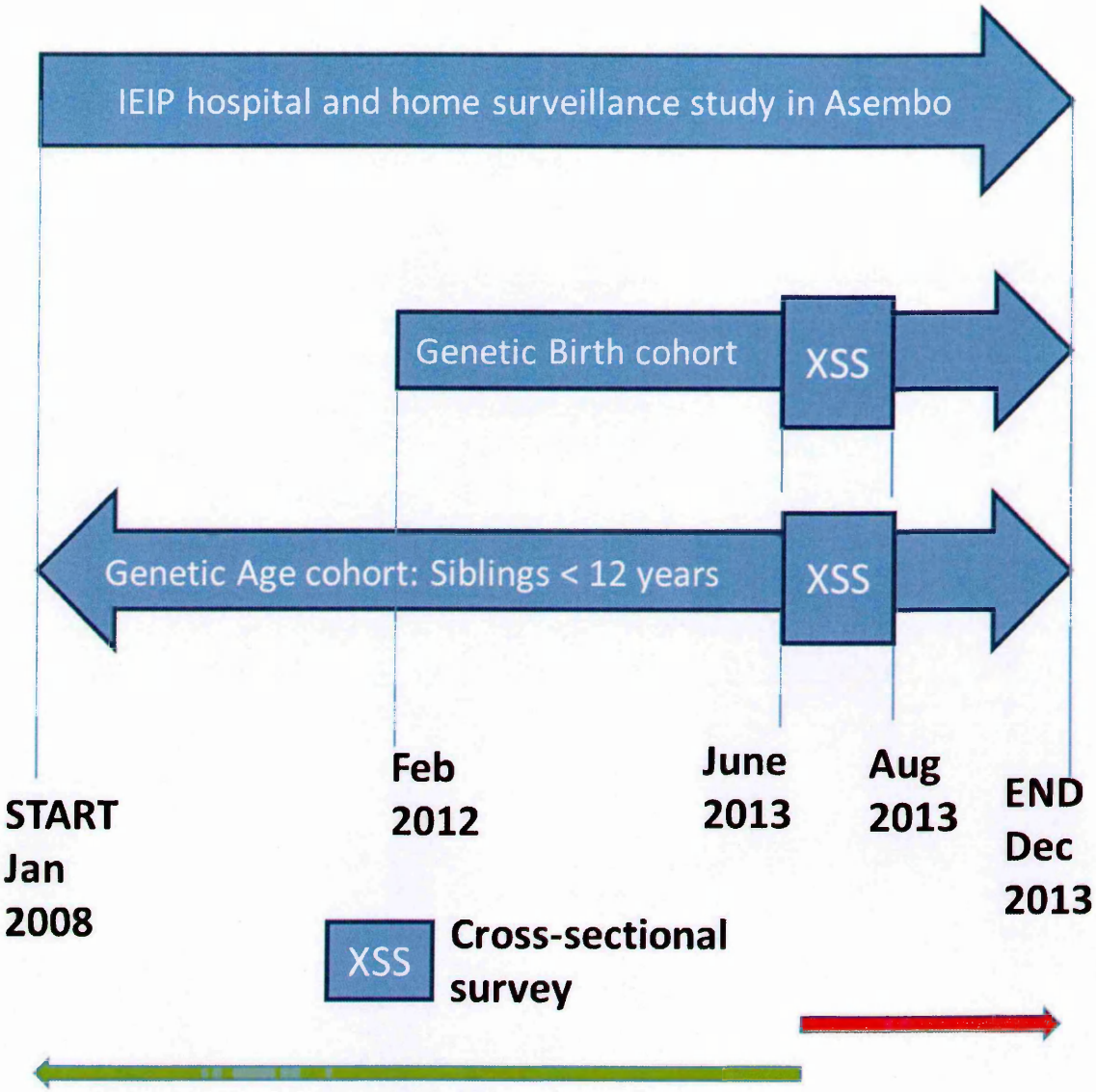
(iii) Follow up for malarial and other childhood disease events.

(iv) Analysis for the effect of genotype on the incidence of disease events.

Since a parasitological case definition of clinical malaria was required, a cross sectional survey involving members of both cohorts was conducted in June-August 2013.

The design of my study is summarised in Figure 2.3 below.

Figure 2. 3 Schematic summary of the study design.



Schematic representation of the study design. The sampling frame was the emerging infections program hospital and surveillance study. From among members of the hospital and home surveillance study, a genetic birth cohort was recruited prospectively from February 2012 to December 2013. A genetic age cohort consisting of siblings of the birth cohort members was recruited in June 2013 to August 2013. The age cohort had two components of surveillance information: a prospective component (red forward arrow) and a retrospective component (green backward arrow). A nested cross sectional survey was conducted among birth cohort and sibling members between June-August 2013.

The major programme was the IEIP study as described in section 2.3 above. The genetic cohort was nested within this programme. It consisted of members followed since birth from February 2012 to December 2013, and a wider variety of ages, termed an age cohort, which was recruited from among the members of the IEIP health and

morbidity surveillance during a cross-sectional survey conducted between June-August 2013. In effect, the data for this study came from the morbidity surveillance, the prospective birth cohort, the cross-sectional survey and the retrospective and prospective age cohort.

The availability of a population-based health and morbidity surveillance system in a population with stable year-round malaria transmission provided me with an excellent opportunity to study the effect of human genetic factors on the incidence of malaria and common childhood illnesses.

2. 4.1 Preparation for the study

Before the study could be initiated, I held a number of meetings to assess the feasibility of the study in the area. These included securing support for the study within the HDSS and enlisting the support of key investigators who were already participating in the health and morbidity surveillance study. These discussions allowed me to develop a protocol detailing the study including the background, aims, study objectives, justification, required budgetary resources, coordination plans, data management, genotyping, data analysis and a publication plan. I then developed a protocol for submission to relevant Scientific and Ethical committees. I was primarily responsible for correspondence between the investigation team and the Scientific and Ethical Committees - copies of these communications are presented in Appendix V.

A series of meetings were held with influential community members including political and religious leaders and workers within the community as follows: meeting the chiefs and provincial administration, Community Advisory Board (CAB), opinion leaders, village reporters and community interviewers. During these meetings, I

presented the aims of the study, the participation of the community, benefits and risks of the study. I received feedback from community members on whether they thought the study would be beneficial to the community and suggestions on how best to conduct the study to mutually benefit the study team and the community.

The study protocol was approved by the KEMRI Scientific Steering Committee in October 2011. Study staff with desired skills in informed consenting, sample collection, laboratory sample processing and reporting were recruited. A number of staff were shared with the main morbidity study, in which case a time sharing agreement was entered into to allow both this genetic study and the main cohort study to proceed concurrently. I trained the study team members according to the protocol and sub protocols regarding the particular procedures in which they would be involved.

Consent forms explaining the purpose of the study, the risks and benefits were prepared in English and then translated into the local languages, Dholuo and Kiswahili. A sample is found in Appendix I. Having prepared the site for the study, recruitment began in February 2012. The major study procedures, including recruitment, genotyping, follow-up and analysis are explained in further detail below.

2.4.2 Recruitment

2.4.2.1 Prospective birth cohort

The birth cohort study targeted all children born within the health and morbidity surveillance area of the Asembo HDSS between 22nd February 2012 and 31st December 2013. The mothers or guardians of these children were approached with a view to recruitment into the study. The objective of the study was explained and where

families were willing informed consent was sought. Eligibility was determined according to the following inclusion and exclusion criteria.

(a) Inclusion criteria

The study included all children who fulfilled the following inclusion criteria:

(i) All children born to mothers participating in the HDSS and IEIP population study.

(ii) Born between 22nd February 2012 and 31st December 2013.

(iii) Less than 6 months old at the time of recruitment.

(iv) Consent obtained from one or both parents or a legal guardian.

(b) Exclusion criteria

(i) Refusal to participate in the study.

(ii) Those who had not enrolled in the IEIP/HDSS study.

(iii) Born before the commencement of the study.

(iv) Born before 22nd February 2012 or after 31st December 2013.

2.4.2.2 Age cohort of siblings

The effect of age on the incidence of malaria was also a key factor under investigation. While it would be possible to capture age-specific events during the first 2 years of life within the birth cohort, this cohort alone would not have allowed me to investigate the genotype-specific incidence of malaria over a wider age range. To complement age effects investigated in the birth cohort, an age cohort covering a wider age range was recruited. Since the age-specific incidence of malaria is a proxy for naturally

acquired immunity (NAI), the upper age limit of 12 years was chosen as an approximate age by which NAI would be presumed to have developed.

I considered a number of approaches to the recruitment of an age cohort covering the age range 0-12 years, including:

- Inviting all children in the HDSS <12 years of age.
- Recruitment of older children for each birth cohort member recruited, matched by area.
- Recruitment of siblings of birth cohort members under the age of 12 years.

The approach chosen would have to be the best for effective control of confounding. In genetic predisposition analyses, confounding would come from two sources; environmental factors and genetic factors. The best method would be one that minimizes the confounding which arises from these sources. Environmental factors may best be controlled by grouping participants who share a common environment in a scale as small as possible. The unit of analysis in HDSS settings is commonly the household, where members share the same socio economic status including sanitation facilities, income and food among other factors, which are likely to be related with the risk of illness. Genetic factors may be controlled for in the analysis if a measure of genetic relatedness among individuals is used. A random sample drawn from the study area would have related individuals by chance, and getting measures of genetic relation between individuals would be a complex task. However, it is known that siblings would share approximately 1/2 of their genes, first cousins about 1/4 of their genes, second cousins 1/8 and so on. Since the study was conducted in a small area (within 5km radius of Lwak hospital) of a typical rural setting where related individuals are found within short distances, failure to adjust for genetic relatedness would potentially

overestimate or underestimate effect sizes seen. From the foregoing, a sibling cohort would be the best approach because it allows for effective control for environmental confounding by having members from the same household, and control for genetic confounding since relationships between siblings would be explicitly captured at recruitment.

This sibling recruitment approach also had practical benefits in the field as it was easier to recruit siblings since their parents had been taken through the study procedures and were more likely to welcome the study team to conduct sibling sampling. In addition, overall, pooling of the samples of children from the younger ages and older siblings would allow the final cohort to attain the power required to detect small effects of genes on malaria incidence. The sampling frame for the sibling cohort was the same as that for the cohort studies, i.e. the Asembo area population within a 5km radius of Lwak hospital and who are already enrolled in the PBIDS study.

Sibling recruitment

Siblings of birth cohort children were recruited during the June-August malaria transmission season of 2013. This was done in the context of malaria cross sectional survey. The same inclusion and exclusion criteria as those for recruiting the birth cohort members as described above were applied for the siblings with the major difference being that the siblings had to be 12 years of age or less at the time of cross sectional survey. Genotyping and disease surveillance follow up was also done for the siblings. The cross-sectional survey is described in more detail below.

2.5. Cross-sectional genetic and malaria survey

2.5.1 Justification

As described in section 1.2.4.8 above, one of my principal aims in this study was to investigate the age- and genotype-specific incidence of malaria among children with a range of inherited red blood cell disorders through a cohort study of children living within the Asembo area. A cross-sectional survey was complementary to these objectives, as it would (i) show the parasite prevalence in the area and (ii) assist in determining the case definition of malaria through a comparison of symptom, fever survey and parasite densities. I therefore conducted a cross-sectional survey of all enrolled and genotyped cohort members to investigate the prevalence and density of malaria parasites along with clinical data on fever.

2.5.2 Conduct of the cross sectional survey: June-August 2013

Scientific and ethical permission were sought from the KEMRI Scientific Steering Committee (SSC) and Ethics Review Committee (ERC) respectively. Once these were secured, I developed a community entry strategy through meetings with the provincial administration and chiefs, followed by meetings with religious leaders and opinion leaders, the community advisory board and finally some of the study workers including village reporters and community interviewers.

Parents of all eligible children were approached in their homes by village reporters and community interviewers and given basic information concerning the study. They were also given invitation cards for each eligible child to come to Lwak hospital on particular days within the survey period. Heeding to the invitation was voluntary. On the designated days between June and August 2013, parents accompanied their

children to Lwak hospital to learn more about the study and if convinced of its' benefits, enrol for the study. On arrival at Lwak hospital, study staff confirmed the eligibility of study participants by checking against a register of all targeted index children and their siblings. Among the eligible members, study staff explained the rationale of the study, i.e. to investigate the link between malaria infection and a range of specific genetic variants.

Each parent who agreed to participate in the study gave written informed consent for participation in the study by signing the informed consent forms. After consenting, parents of the children answered a questionnaire on malaria risk factors separately for each child to distinguish unique risk-factors which may predispose particular children to malaria more than others. Temperature measurements were then taken for each child to record the temperature on the day of the survey, compared to reports of fever during the week. Thermoscan^{TM} digital thermometers were used to record axillary temperatures.

A blood sample was collected using a finger prick procedure for children >1 year of age and heel prick for those children <1 year old. Thick and thin malaria blood films were made for diagnosis of malaria by microscopy. The malaria slides were then left for air-drying before staining with 10% Giemsa stain. Due to logistical constraints, slides were not read immediately but stored appropriately for speciation and parasite density readings. Slides were examined at x1000 magnification for asexual *P. falciparum* parasites. 100 fields were examined before slides could be considered negative.

During the cross sectional survey, the diagnostic test used for case management was a rapid diagnostic test (RDT) Carestart™ *Pf HRP2* (American Access Bio Company) kit. In addition, dry blood spots (DBS) were collected as a back up to the whole blood collected from each individual.

All children found with RDT positive results were treated with a combination of Artemether and Lumefantrine (AL), the first line treatment for uncomplicated malaria in Kenya. Dosage of AL was based on the weight of the child for maximum effectiveness. Weighing of children was done to confirm the weight of the child before administration of AL. In accordance with World Health Organization and National guidelines, the administration of the AL was done in Lwak hospital and the child observed to confirm that they took the first dose. For those who vomited within 30 minutes, a repeat administration was done to ensure the correct dose of the drug was taken. Their parents were also advised to monitor them for vomiting of dosages taken at home, in which case they would re-administer the drug.

During the cross sectional survey only siblings had an EDTA sample of about 200µl collected from the finger prick for genotyping for red cell genetics and other malaria susceptibility genes. EDTA samples were not collected for index children as these had already been collected earlier for genotyping and ABO blood grouping by serology. Serological typing for ABO blood groups was additionally done for siblings. During the survey period, the samples were kept in -20^{0}C freezers awaiting full genotyping later in the study before being transported to the KEMRI/CDC Kisian laboratories. Stained slides were kept in slide boxes pending later reading at the KEMRI Centre for Geographic Medicine Research-Coast (CGMRC) in Kilifi.

2.6 Surveillance for clinical malaria episodes in the longitudinal cohort study

Surveillance for clinical malaria episodes during the cohort follow-up was done through the IEIP system as described above in section 2.3.2.

2.6.1 Case definition of malaria in the cohort study

Malaria may either be uncomplicated or complicated. Uncomplicated malaria is defined as symptomatic infection with malaria parasitaemia but without signs of severity. Severity may or may not include vital organ dysfunction (WHO, 2010). Complicated malaria, also described as severe malaria refers to syndromes of malaria that cannot be explained by other aetiologies in a child positive for malaria slides. These include severe anaemia, respiratory distress and cerebral malaria (WHO, 2000), as described in section 1.2.3.2.2 above. Since fevers are common in endemic areas, parasite density above a certain threshold would be required for a more specific diagnosis of malaria. During the cohort follow-up in Lwak hospital, data were available regarding Giemsa-stained blood-slide positivity, inability to breastfeed, lethargy and convulsion, difficulty in breathing and data on a comma scale, the AVPU response scale, in which the letters refer to whether the patient is Alert, responds to Voice, responds to Pain or is Unconscious. From these data, a pragmatic case definition of malaria was adopted as follows.

Uncomplicated malaria was defined as fever in the presence of a asexual parasite positive Giemsa-stained slide, in a child who was not admitted to hospital, was alert (A) or responding to voice (V) on the AVPU scale, who lacked any of the severe symptoms listed above and did not have an alternative bacterial cause for their fever.

Complicated malaria was defined as fever in the presence of a malaria-positive blood film in a child admitted to hospital or referred for care who also presented with severe disease symptoms such as inability to breastfeed, convulsions, difficulty in breathing or unconsciousness (U) on the AVPU scale. Table 2.2 below gives the definition for uncomplicated and severe malaria that were used in the current study, as derived from the WHO definitions.

Table 2. 2 Data available from the Lwak hospital IEIP clinic surveillance corresponding to WHO definition of uncomplicated and severe malaria.

Case definition used	**WHO definition uncomplicated malaria (WHO, 2010)**	**WHO severe malaria definition (World Health Organization, 2000)**	**Data available in Lwak hospital questionnaire** (ref Appendix)
Fever or history of fever in last 48 hours	Fever	Symptoms include headache, lassitude, fatigue, abdominal discomfort, and muscle and joint aches, usually followed by fever, chills, perspiration, anorexia, vomiting and worsening malaise (WHO, 2010)	Q. 3.1.0 presenting with fever Q 5.4 history of fever Q 6.0 Case definition of fever met
Blood film positive for malaria parasites	Blood film positive on 10% Giemsa stain	*P. falciparum* asexual parasitaemia	6.3b Giemsa stain results (Positive/Negative)
	Parasite species identification		6.3c Parasite species?
	Lacks any of the severe signs of vital organ function below	Has any of the severe signs identified below	See below for questionnaire indications
Any one or more of these severe malaria syndromes		Prostration	Q. 5.1a Able to breastfeed (Yes/No)
		Impaired consciousness	Q 5.13 Unconscious (Yes/No)
		Respiratory distress	Q 5.2 Difficult breathing (Yes/No)
		Multiple convulsions	Q5.14 Convulsion or History of convulsion (Yes/No)
		Severe anaemia	5.33 Clinical signs of anaemia e.g. paleness of palms, nail bed or conjunctiva (Yes/No) 6.1 Haemoglobin g/dl

2.6.2 Definition of other childhood infectious disease syndromes

Definitions of syndromes used were adapted from the IEIP as follows.

Respiratory illness was defined for those <5 years as cough or difficulty breathing and one of the following: convulsions, unable to drink fluids or unable to breastfeed, lethargic, chest in drawing, vomiting everything, stridor, oxygen saturation <90%; and for persons ≥5 years old as cough or difficulty breathing or chest pain and one of the following: temperature $\geq 38^0$C and oxygen saturation <90%.

Non malarial fever: documented axillary temperature $\geq 38^0$C without an obvious cause, defined as cough, difficulty breathing, chest pain, signs of meningitis, or bloody diarrhoea. Presenting with symptom of fever together with no diagnosis of malaria and a slide negative for malaria parasites by the Giemsa stain.

Diarrhoea or gastroenteritis: Those with a diagnosis or discharge diagnosis of diarrhoea or gastroenteritis or dysentery. Diarrhoea was defined as at least 3 looser than normal stools reported in a 24 hour period (Feikin *et al.*, 2011).

2.7 Host genotyping

Recruitment blood samples from all cohort members were typed for a range of candidate malaria-associated genes by a variety of methods.

Haemoglobin phenotyping

Haemoglobin (Hb) typing for normal (HbAA), carrier (HbAS) or sickle cell anaemia (HbSS) status was done by cellulose acetate electrophoresis using the Helena systems kit according to the manufacturer's instructions (Laboratories). Briefly, electrophoresis consisted of a lane with a mix of controls for normal haemoglobin (HbA), foetal haemoglobin (HbF) sickle haemoglobin (HbS) and haemoglobin C (HbC), which was run alongside sample lanes. On completion of the electrophoresis run, sample bands corresponding to the control band positions for haemoglobin A, F, S, C were interpreted as either A, F, S or C. It was then possible to have phenotypes as AA, FF, SS, CC, AF, AS, AC, FS, FC, SC as follows:

Table 2. 3 Interpretation of results for haemoglobin electrophoresis.

Sample band corresponding to control band at position	**Phenotype**
A only	Hb AA
F only	Hb FF
S only	Hb SS
C only	Hb CC
A and F	Hb AF
A and S	Hb AS
A and C	Hb AC
F and S	Hb FS
F and C	Hb FC
S and C	Hb SC

ABO phenotyping

To type each sample for ABO blood groups, 3 separate drops of whole blood were dropped onto a microscopy slide. Using a dropper, drops of antisera for blood group A, B and Rhesus blood group were used to spot the blood drops in turn. Blood was mixed with antisera using a rod and the slides were checked for agglutination and interpreted as per table 2.4 below.

Table 2.4. Serological phenotyping for the ABO blood groups.

Reaction with antisera	Interpretation
A	Blood group A
B	Blood group B
A and B	Blood group AB
D	Rhesus positive +
No reaction with either anti A or anti B	Blood group O

It was then possible to discern Rhesus positive blood groups A+, B+, AB+, O+ or Rhesus negative blood groups A-, B-, AB- and O-.

2.6.1 DNA-based genotyping for other candidate genes

Molecular DNA based methods were used to confirm phenotypes for haemoglobin electrophoresis and ABO blood groups. CR1, G6PD deficiency and 59 additional SNPs chosen from a review of literature for malaria candidate genes by the malaria genetic epidemiology network, MalariaGEN (http://www.malariagen.net/), were also

typed using molecular methods. These SNPs, which are in immunity and inflammation genes, cytokine genes, red blood cell conformation and metabolism genes among others, are listed in full in Appendix III.

2.6.1.1 DNA extraction

With the help of colleagues in the laboratories at the KEMRI-CDC Programme in Kisumu, and the KEMRI-Wellcome Trust Programme in Kilifi, I extracted genomic DNA from whole blood using the QIAamp DNA Blood Mini Kit (Qiagen) according to manufacturer's instructions. DBS samples from the same individuals which had been collected were used as a back-up sample to allow re-extraction in case the whole blood sample was damaged. The whole blood sample extraction is briefly described. 20µL of proteinase K was pipetted into a micro centrifuge tube. 200µL of whole blood was then added to the tube. Where the sample was less than 200µL it was topped up to a volume of 200µL using phosphate buffered saline (PBS). 200µL of buffer AL was then added to the sample, and pulse vortexing used to mix the sample for 15 seconds. The mixture was incubated at 56^0C for 10 minutes to allow lysis of cells to release DNA into the mixture. The tube was then centrifuged to remove drops from the inside of the lid. DNA was precipitated by adding 200µL of ethanol (96-100%) to the micro-centrifuge tube. Pulse vortexing was done for 15 seconds to allow thorough precipitation and the tube was spun briefly to remove drops from the inside of the lid. The mixture containing the precipitate was then pipetted onto a spin column with a collection tube underneath. It was centrifuged at 6000g (8000 revolutions per minute (rpm)) for 1 minute to allow the DNA to be trapped in the spin column and the rest of the mixture to be collected into the collection tube. The collection tube containing the non-DNA mixture was then discarded and the spin column washed twice with wash buffer 1 (AW1) and wash buffer 2 (AW2) respectively to further purify the DNA.

Wash 1 was done by adding the wash buffer (AW1) to the spin column and centrifuging at 8000rpm for 1 minute before placing the spin column in a new collection tube. Wash 2 was done by adding the wash buffer (AW2) to the spin column and centrifuging at 20,000g (14,000rpm) for 3 minutes. After the wash steps above, the spin column was placed onto a 1.5ml micro-centrifuge tube. DNA was eluted from the spin column by adding 200μL of the elution buffer (Buffer AE, or distilled water). The spin column was incubated for 1 minute at room temperature (23-25^{0}C) and then centrifuged at 6000g (8000rpm) for 1 minute. The spin column was discarded and the eluted DNA in the micro-centrifuge tube was then stored for further DNA analysis. In order to type for SNPs using multiplex technology, DNA was shipped frozen to Oxford, where the remaining genotyping was conducted by collaborators.

2.6.1.2 Typing for the α^+thalassaemia deletion

The 3.7kb gene deletion in the α-globin gene resulting in α^+thalassaemia was typed separately by a specific PCR reaction (Chong *et al.*, 2000). A Qiagen PCR kit was used to make a PCR master mix volume of 10.05μL per reaction, consisting of 2μL of RNase free water, 2μL of 5X Q-Solution, 3 μL of 2X fast cycling PCR master mix, 1 μL of 10X Coral load cycling dye and α^+thalassaemia specific primers (10 μM) as follows: 0.375μL of Primer 376, 0.375 μL of Primer 377, 0.1 μL of primer 378, 0.6 μL of LIS1-F and 0.1 μL of LIS1-R.

The primer sequences were as below:

Primer LIS1 F	(GTCGTCACTGGCAGCGTAGATC)
Primer LIS1-R	(GATTCCAGGTTGTAGACGGACTG)
376 (α2F, -3.7F)	(CCCCTCGCCAAGTCCACCC)

378 (α2R) (AGACCAGGAAGGGCCGGTG)

377 (-3.7R) (AAAGCACTCTAGGGTCCAGCG)

The master mix was plated into wells containing 2μL of control (homozygous α^{+}thalassaemia control, heterozygous control, normal control and a no template control) or sample DNA. The plate was carefully capped and cycling was done according to the following profile: 95^0C 5:00 min, 30 cycles of (97^0C 0:45 min, 60^0C 1:15 min, 72^0C 2:30 min) and a final extension of 72^0C for 5:00 min. After PCR, 2 μL of the amplicon was run on a 1% agarose gel at 95V for 2hours and 30minutes. The gel was then exposed to an ultraviolet ray trans-illuminator (GelDoc), where a single 1.8Kb band was interpreted as αα/αα (normal), a 2.1 Kb band was interpreted as $^{-3.7}\alpha/^{-3.7}\alpha$ (homozygote) and two bands consisting of 2.1Kb and 1.8Kb was interpreted as αα $/^{-3.7}\alpha$ (heterozygote).

2.6.1.3 Sequenom Mass Array™ genotyping

At the Wellcome Trust Centre for Human Genetics (WTCHG), Sequenom Mass Array™ genotyping was used to confirm genotypes for sickle and ABO blood groups and to assay for a total of 65 SNPs in malaria candidate genes as follows. Genomic DNA samples were used to conduct whole genome amplification using the primer extension pre-amplification technique as described by Zhang and colleagues (Zhang *et al.*, 1992). The whole genome amplicons were then genotyped using the Sequenom Mass Array™ genotyping platform (P. Ross *et al.*, 1998; Wilson *et al.*, 2005). The platform consists of specific PCR amplification using sets of primers, which are able to distinguish between different SNPs since the size or base composition of the amplicon is different according to the SNP present. Mass spectrometry of the PCR

products allows the resolution of the mass of the PCR product, decoding the masses of particular bases. Each mass window is unique to a particular primer's target amplicon. It is therefore possible to have a multiplex system where several targets are amplified as long as the mass windows of their amplicons are different. Different mass windows are assured by two methods, either having different lengths of amplicons or using different base composition of amplicons. A more detailed description of the method is included in Appendix IV.

2.7 Statistical analyses

The incidence of malaria over time was assumed to follow a Poisson distribution, as previously published (Williams, Mwangi, Roberts, *et al.*, 2005). A univariate model predicting malaria incidence by genotype was fitted for the 65 SNPs assayed and for α^{+}thalassaemia genotypes. All genotypes which had $P<0.1$ in the univariate model were fitted in a multivariate model to predict malaria incidence. The multivariate model in addition to adjusting for effects of other genes adjusted for covariates of malaria including age, gender, socio-economic status (SES), mothers/household head's education level, year, season and distance to Lwak hospital. The estimates were also adjusted for clustering by household and genetic relationship using sib-ship information. Estimates were further refined using the sandwich estimator (Armitage P, 2001), which inflates confidence intervals to account for correlation between observations events on the same child.

Variable descriptions

Gender: A categorical variable consisting of 2 levels, male or female.

Age: A continuous variable, converted to categorical age bands of 3 months up to 2 years, and 2 year age bands up to 10 years of age.

Mothers' or household head's education level: This was a categorical variable which could be; no education, primary, secondary or post-secondary.

Distance to Lwak hospital: The straight line distance from the participants household to Lwak hospital, calculated using geographical coordinates of the household and the hospital.

Socio-economic status: Defined in wealth quintiles with 1 being the most poor and 5 being the least poor.

Season: A binary categorical variable, with high season being during the long rains and low season being during the dry season and short rains.

Testing for interactions: I used the likelihood ratio method to compare Poisson models with and without interaction terms to test for interactions between exposure variables including genetic polymorphisms. Where multiple SNPs were tested (see Chapter 5), the p-value for significance was adjusted using the Bonferroni correction method.

Model of inheritance

An additive model of inheritance was assumed for these analyses. This model performs well in detecting effects of SNPs that are inherited in an additive or co-dominant fashion. However, this model does not maximize power for those genes where recessive effects contribute to the disease being studied (Lettre *et al.*, 2007).

CHAPTER 3: The epidemiology of malaria and other childhood infectious syndromes in Asembo

In this chapter I present a summary of the epidemiology of malaria and of other common diseases of childhood within the context of my cohort study. Epidemiological risk factors predicting malaria incidence were investigated using two approaches, a longitudinal approach and a cross sectional survey. The risk factors investigated included demographic, seasonal, socio-economic and health seeking behaviour. In the longitudinal approach, since the disease rates presented here were calculated from data collected in Lwak hospital, the rate of health seeking for potential malaria fevers was investigated and the results are presented. In addition, the parasitological case definition for malaria was explored through a cross-sectional survey of malaria fevers confirmed by blood films. The rates of infectious syndromes such as respiratory tract infections, diarrhoea and non-malarial fevers are presented in comparison with the rate of malaria so as to position the malaria burden in relation to other causes of childhood morbidity in the area.

3.1 Objectives

The specific objectives of this chapter are as follows:

1. To investigate the epidemiology of malaria in children living in the IEIP area through a longitudinal cohort study.

2. To further investigate the epidemiology of malaria in the same area through a cross sectional study.

3. To investigate the parasite density threshold associated with malaria fever.

4. To compare the incidence of malaria with that of other infectious syndromes detected through the IEIP system.

Two approaches were used to define the epidemiological risk factors for malaria; a cohort approach and a cross sectional survey approach. The cohort approach was better suited to evaluate time changing variables such as season, calendar year and age, whereas the malaria cross sectional survey approach was useful in delineating specific factors related to malaria including bed net use, anti-malarial use and level of education of the household head. The following sections give determinants of malaria in both cohort and cross sectional designs.

3.2 The epidemiology of malaria in the genetic cohort: Longitudinal study

3.2.1 Incidence of uncomplicated malaria within the cohort

Over a period of 6 years (January 2008- December 2013), 2431 disease episodes of malaria were recorded in 4936.3 child years of follow-up, resulting in a rate of 0.49 (95% CI 0.47-0.51) malaria episodes/child year of follow up (cyfu). The cohort consisted of 1462 children, index children were 659 and siblings were 803. The calculation of person time discounted the loss of follow up time due to being away from the study area. Two considerations were used to determine time away: residency in the area and the time to malaria symptoms once infected. Residency was defined as being present during HDSS census rounds every 4 months. Only those who had an exit event were noted, as such the absence of an exit event from the HDSS indicated the presence of the individual during the census round. The second consideration; time

to malaria symptoms while infected, would be the pre-patent (incubation) periods for the species of plasmodia prevalent in the study area. The periods for the two species present are; 7 days for *P. falciparum* and 30 days for *P. malariae*. The time period of 7 days was considered as too sensitive and may have resulted in an overestimation of the loss to follow up time, and therefore a 30 day period was chosen as defining time away from the study area.

Since the current study had access to both HDSS follow-up data (every 4 months) and home and morbidity surveillance data (every 2 weeks), a more stringent definition of follow-up time could be made from these two data sources. For HDSS data, time away was defined as the difference between an exit event and an entry HDSS event. The same individuals were further assessed for time away using household morbidity surveillance (HMS) data. For HMS data, time away was defined as a difference between consecutive HMS interview dates of more than 30 days, considering the incubation period as described above.

Among the birth cohort members, 45 individuals had 5.7 years discounted which consisted of a difference between consecutive HMS visits. Among the sibling cohort members, 699 individuals had 224 years discounted consisting of 43.6 years between an exit and entry event and 180.2 years due to differences of more than 30 days between consecutive HMS visits.

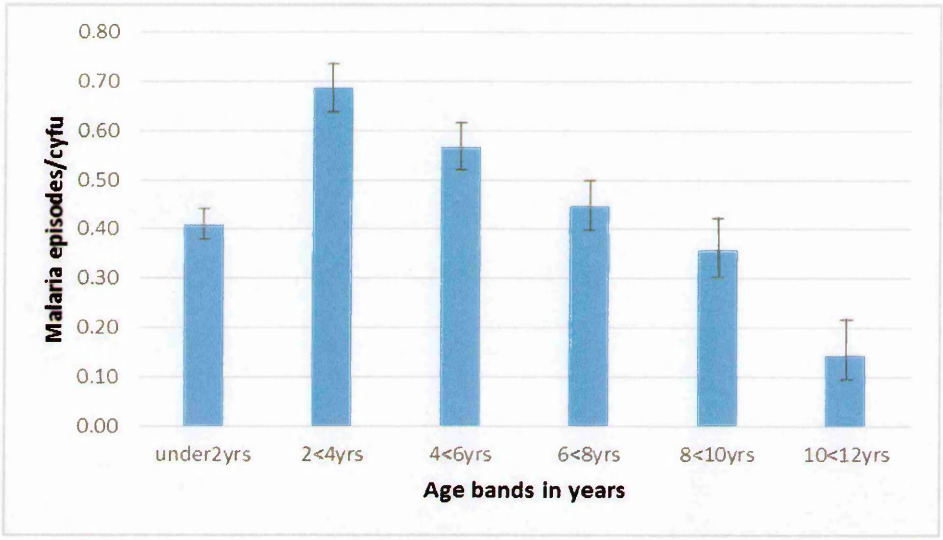
No follow-up time could occur after death, so the data were checked to ensure this was the case for all the deaths recorded. There were 18 deaths in the birth cohort and only 1 death recorded in the age cohort. The lower number of deaths in the age cohort

indicates a selection bias, the survivor effect, since the sibling death events could not be observed among the sibling cohort members retrospectively. In summary after consideration of all HDSS events, including deaths, outmigration, exit and HMS loss of follow-up time, the total person time accumulated was 4936.3 years. Factors likely to affect the rate detected include health-seeking behaviour for disease symptoms and the case definitions used for malaria in the hospital. Health seeking will be addressed in the discussion section. Malaria illness was defined as described in Table 2.2 in the hospital. In brief, a malaria episode was defined as a positive slide in the presence of fever, with accompanying symptoms such as muscle and joint aches, malaise and headache. This rate is lower than the rates reported in the area by the RTS, S malaria vaccine trials (Agnandji *et al.*, 2012), of about 4 episodes per child per year. Personal communication with the Siaya study team was used to confirm the site specific rates of malaria recorded during the RTS, S study. Differences in follow up procedures between the current study and the RTS, S study may account for differences in the rates of malaria recorded. In the RTS, S trial parents or guardians were encouraged were informed to look out for signs of illness and transportation was facilitated. In addition, all participants presenting to an RTS, S trial facility with a reported or documented fever during the previous 24 hours had evaluations done for malaria. In the current study, parents reported to health facilities at their discretion and transportation was not provided. In addition, not all participants with 24 hour reported fever had evaluations done for malaria.

The incidence of malaria by age

The incidence of malaria was low in children <2 years, increased thereafter through to a peak at 2-4 years then showed a reduction in all the older age groups (Figure 3.1).

Figure 3. 1 The incidence of malaria within the cohort by age.



Children were followed from 2008-2013. Incidence data was calculated in age strata, therefore one child could contribute data to several age strata during the study period. The number of children contributing data per age strata are as follows: <2 years N=1454 cyfu=1691.7; 2<4 years N=845 cyfu=1095.1; 4<6 years N=697 cyfu=915.2, 6<8 years N=480 cyfu=692.2; 8<10 N=289, cyfu=388.4; 10<12 N=151 cyfu=153.7 The blue bars represent the crude incidence, error bars represent the 95% confidence interval.

The incidence of malaria by calendar year

Incidences of malaria by year are summarised in Table 3.1 below. The adjusted incidence of malaria was similar during 2008 and 2009, after which a marked increase was seen in 2010. After this, a significant decline was observed in 2011, with similar rates of malaria for the years 2011, 2012 and 2013 being observed in the cohort.

The effect of season

The seasonal indicator most relevant to malaria transmission in this area is the amount of rainfall since it affects the abundance of malaria vector breeding sites. The major rainy season in Asembo occurs between April-July while a shorter rainy season may occur in November or December. To study the effect of season, therefore, the months

throughout the year were divided into two: April to September, which includes the long rains season and October to March, which includes the short rains season. The incidence rates of malaria in these two seasons were compared. Higher incidence of malaria was observed in the long rains season. Using season as a dependent variable, a Poisson regression model was fitted to the data to predict malaria incidence, with age, year and distance to hospital as covariates. The short rainy season was used as the reference group in the analysis. The adjusted incidence rate ratio for malaria incidence in the rainy season was 1.24 (95% CI 1.14-1.35; P<0.0001). Season is therefore an important variable to adjust for in further investigations of malaria by risk factors such as genetics.

Effect of gender, mother's education and socioeconomic status (SES)

Gender significantly affected the likelihood of having a malaria episode. Compared to boys, girls had a higher incidence of malaria, 1.17 (95% CI 1.07-1.27; P=0.01). This may be due to societal values of boys and higher health seeking behaviour for them. Mother's education also had a significant effect. Compared to primary educated mothers, secondary educated mother's children had a reduced incidence of malaria, 0.64 (95% CI 0.56-0.73; P<0.0001). There was no significant effect of tertiary education, although the trend showed a higher incidence of malaria. This could be due to the low number (n=13) of tertiary educated mothers in the sample.

Regarding socioeconomic status, less poor SES had a higher incidence of hospital diagnosed malaria compared to the poorest (see table 3.1). This finding appears counter-intuitive since malaria is expected to affect the poor the most. However, since

the current study was a hospital based study, this may reflect better health seeking among higher socioeconomic classes.

The effects of age, year, season, gender, distance to hospital, mother's education, and socioeconomic status on longitudinal incidence of malaria are summarized in Table 3.1

Table 3. 1 Effect of epidemiological factors identified longitudinally on the incidence of clinical malaria.

	No. of individuals	cyfu	Disease events	Incidence rate (events/cyfu)	Crude IRR (95% CI)	p-value*	Adjusted IRR** (95% CI)	p-value**
Overall rates	1462	4936.3	2431	0.49 (0.47-0.51)	NA	NA	NA	NA
Year								
2008	466	353.32	173	0.49 (0.42-0.57)	1	NA	1	NA
2009	584	599.74	279	0.47 (0.41-0.52)	0.95 (0.79-1.15)	0.6	0.93 (0.76-1.14)	0.49
2010	724	712.92	540	0.76 (0.70-0.82)	1.55 (1.30-1.84)	<0.0001	1.66 (1.38-1.99)	<0.0001
2011	789	850.95	430	0.51 (0.46-0.56)	1.03 (0.87-1.23)	0.73	1 (0.83-1.21)	1
2012	1219	975.14	423	0.43 (0.39-0.48)	0.89 (0.74-1.06)	0.18	0.86 (0.71-1.04)	0.11
2013	1448	1444.23	586	0.41 (0.37-0.44)	0.83 (0.70-0.98)	0.03	0.88 (0.74-1.06)	0.18
Season								
Oct-Mar (Short rains)	1415	2367	1028	0.43 (0.41-0.46)	1	NA	1	NA
April-Sept (Long rains)	1432	2369	1403	0.55 (0.52-0.58)	1.26 (1.16-1.36)	<0.0001	1.24 (1.14-1.35)	< 0.0001
Age								
0-2	1454	1691.73	692	0.41 (0.38-0.44)	1	NA	1	NA
2-4	845	1095.06	751	0.69 (0.64-0.74)	1.68 (1.51-1.86)	< 0.0001	1.43 (1.28-1.60)	< 0.0001
4-6	697	915.21	518	0.57 (0.52-0.62)	1.38 (1.23-1.55)	< 0.0001	1.22 (1.08-1.38)	< 0.01
6-8	480	692.15	309	0.45 (0.40-0.50)	1.09 (0.95-1.25)	0.2	0.98 (0.85-1.14)	0.83
8-10	289	388.42	139	0.36 (0.30-0.42)	0.87 (0.73-1.05)	0.15	0.83 (0.69-1.01)	0.06
10-12	151	153.72	22	0.14 (0.09-0.22)	0.35 (0.23-0.53)	< 0.0001	0.35 (0.22-0.55)	< 0.0001
Gender								
Boys	729	2359.7	1162	0.49 (0.46-0.52)	1	NA	1	NA
Girls	733	2314.4	1269	0.55 (0.52-0.58)	1.11 (1.03-1.21)	0.01	1.17 (1.07-1.27)	0.01
Effect of distance								
0-2km	182	613.78	639	1.04(0.96-1.13)	1	NA	1	NA

	No. of individuals	cyfu	Disease events	Incidence rate (events/cyfu)	Crude IRR (95% CI)	p-value*	Adjusted IRR** (95% CI)	p-value**
2-4km	602	1935.5	1232	0.64 (0.60-0.67)	0.61 (0.56-0.67)	<0.0001	0.59 (0.54-0.65)	< 0.0001
>4km	546	1823.61	351	0.19 (0.17-0.21)	0.18 (0.16-0.21)	<0.0001	0.18 (0.16-0.20)	< 0.0001
Education								
Primary	1204	4049.16	2037	0.50(0.48-0.53)	1	NA	1	NA
Secondary	265	773.47	331	0.43 (0.38-0.48)	0.85 (0.76-0.96)	0.01	0.64 (0.56-0.73)	<0.0001
Tertiary	13	35.38	8	0.23 (0.11-0.45)	0.45 (0.22-0.90)	0.02	1.19 (0.59-2.39)	0.63
SES								
Most poor	141	409.27	187	0.46(0.40-0.53)	1	NA	1	NA
More poor	271	950.09	497	0.52 (0.48-0.57)	1.14 (0.97-1.35)	0.11	1.36 (1.14-1.63)	<0.001
Poor	324	1108.2	611	0.55 (0.51-0.60)	1.21 (1.02-1.42)	0.02	1.35 (1.14-1.62)	<0.001
Less poor	327	1088.68	584	0.54 (0.49-0.58)	1.17 (1.00-1.38)	0.06	1.27 (1.06-1.52)	0.01
Least poor	418	1342.95	541	0.4 (0.37-0.44)	0.88 (0.75-1.04)	0.14	1.17 (0.98-1.40)	0.09

cyfu, child-years of follow-up; REF, reference group; CI, confidence interval; IRR, incidence rate ratio; NA, not applicable;

* Not adjusted

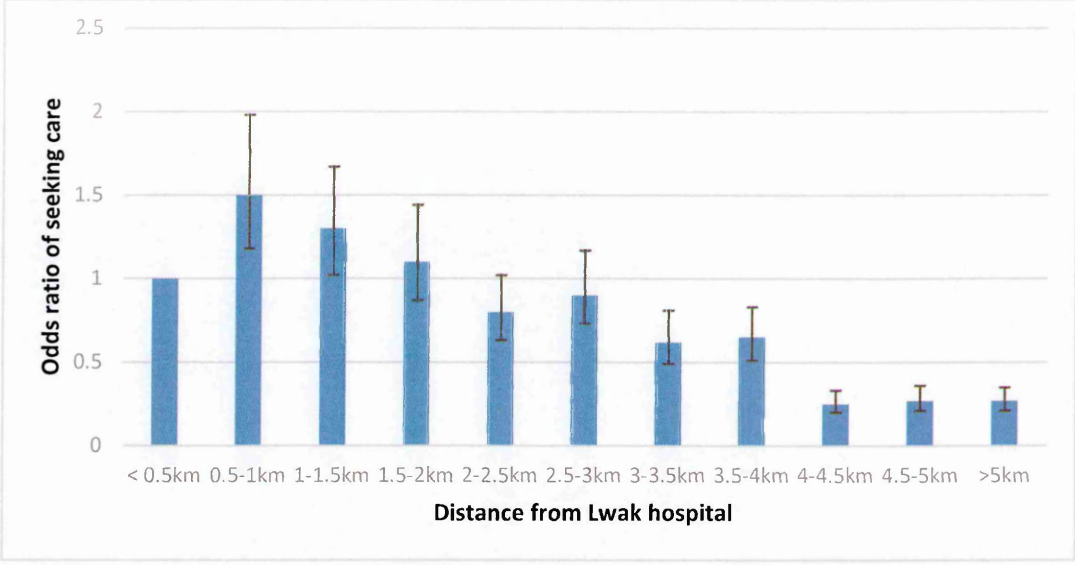
** Adjusted for age, year, season and gender, socio-economic status, distance to hospital and mother's education. Adjustments were done by including these variables in a multivariable model.

Effect of distance to Lwak hospital

Since this study relied on passive case detection, I hypothesized that health seeking behaviour might play a major role in the rates of malaria reported. A previous study in the area (Bigogo *et al.*, 2010) showed that residents sought care at Lwak hospital for only 10% of the fevers reported through the fortnightly home visits. Therefore, it may be expected that the 'true' incidence of malaria that would be captured at home visits might be many times higher than that detected through this surveillance approach. For our study, we observed a malaria incidence of 0.49 episodes per child per year, and the therefore the true community incidence rate might actually be as high as 4-5 episodes per child per year, similar to that estimated in the RTS, S trial (Agnandji *et al.*, 2011) of 4 to 4.5 episodes per child per year.

Self-reported care-seeking at Lwak hospital from the home morbidity surveillance was analysed by a logistic regression to determine the odds of seeking care at the hospital. The results are presented in Figure 3.2 below.

Figure 3. 2 Odds of seeking care for fever stratified by distance from the hospital.



The blue bars represent the crude estimate of the odds ratio for care seeking at the hospital for different distances, while the error bars represent the 95% confidence interval. The reference group is the odds of care seeking for cohort members living within half a kilometre from Lwak hospital (OR=1), compared to individuals living in successive distances of 500m up to 5 kilometres and over 5 kilometres from Lwak hospital.

These results show a significant difference in the odds of seeking care in Lwak between those who are less than 2km from the hospital, those 2-4km from the hospital and those 4km and over from the hospital. Distance is therefore be an important factor to adjust for in studies investigating the incidence of disease within the study framework.

3.2.2 Incidence of severe or complicated malaria

A total of 41 episodes of severe malaria were recorded within the cohort during a follow-up period of 4936.3 years, giving a rate of severe malaria of 8 episodes per 1000 child years of follow-up. As expected the rate of severe malaria was considerably lower than that of uncomplicated disease. Comparison of complicated and uncomplicated disease shows that approximately 2% of disease episodes progressed to become severe. Since

mild malaria cases were underestimated by a factor of 10, severe malaria cases could also have been underestimated by the same factor.

3.3 The epidemiology of malaria in the genetic cohort: cross-sectional survey

A cross-sectional survey was conducted between June 2013 and August 2013 with the aim of gaining an understanding of the probability of having malaria parasites in this community. A questionnaire, which included questions on factors that might possibly affect the probability of having malaria parasites, was designed and administered to the study participants (see Appendix VI).

The list of possible risk factors investigated included:

i) Level of education of household head.

ii) Age of participant.

iii) Socioeconomic indicators of the household such as wall type, roof type and floor type.

iv) Influence of malaria control efforts such as e.g. IRS (indoor residual spraying) and the use of mosquito coils and spray repellent.

v) Bed-net use.

vi) Gender.

vii) Sought care or not.

viii) Anti-malaria drugs taken recently.

Complete parasitological, risk factor and genotype data was collected on a total of 1044 individuals. The risk of malaria was assessed using a malaria case definition of fever in the presence of a positive malaria blood film.

Factors with a significant effect on malaria in the crude analysis included higher level of education of the household head, use of at least one antimalarial and seeking any sort of care. However, in the adjusted analysis only age (older participants at 6-8 years had a higher risk of malaria compared to the youngest group of 0-2 years) and whether the participant sought care remained significant. Other factors were not associated with protection of susceptibility to malaria. These factors included gender, bed-net quality, indoor residual spraying in the last 12 months and repellent use. House type indicators, which were used as a proxy for socio-economic indicators, did not show any effect on any case definition.

Table 3. 2 Effect of risk factors identified at cross-section on odds of having malaria (defined as a positive slide and a fever in the last 24 hours).

	Number of individuals (%)	No. of episodes	p-value*	OR* (95% CI)	Adjusted OR** (95%CI)	p-value**	Type III P-value
Total	1044 (100)	57	NA	NA	NA	NA	
Age							
0-2	287 (27.4)	13	NA	REF	NA	NA	0.03
2-4	192 (18.3)	14	0.2	1.66 (0.76-3.61)	1.56 (0.70-3.50)	0.28	
4-6	176 (16.8)	12	0.29	1.54 (0.69-3.46)	1.56 (0.68-3.57)	0.29	
6-8	168 (16.1	14	0.1	1.91 (0.88-4.18)	2.29 (1.00-5.27)	0.05	
8-10	122 (11.7)	1	0.09	0.17 (0.02-1.35)	0.21 (0.03-1.62)	0.13	
10-12	99 (9.5)	3	0.52	0.66 (0.18-2.36)	0.82 (0.22-2.99)	0.76	
Antimalarial use							
No	909 (87.1)	41	NA	REF	NA	NA	0.58
Yes	135 (12.9)	16	0.006	2.79 (1.50-5.18)	1.27 (0.55-2.95)	0.58	
Education level of household head							
None	53 (5.1)	5	NA	REF	NA	NA	0.08
Primary	709 (68.2)	44	0.38	0.65 (0.25-1.72)	0.92 (0.31-2.76)	0.88	
Secondary	262 (25.2)	8	0.05	0.31 (0.10-1.00)	0.49 (0.14-1.76)	0.28	
Tertiary	16 (1.5)	0	NA	NA	NA	NA	
Sought care for fever							
Yes	236 (24)	26	NA	REF	NA	NA	<0.0001
No	749 (76)	29	< 0.001	0.33 (0.19-0.58)	0.36 (0.16-1.78)	0.01	
Gender							
Boys	550 (52.9)	26	NA	REF	NA	NA	0.29

Girls	490 (47.1)	31	0.25	1.36 (0.80-2.34)	1.27 (0.73-2.24)	0.39	
Bed net quality							
Bad	316 (31.4)	21	NA	REF	NA	NA	-
Good	691 (68.6)	32	0.13	0.64 (0.36-1.14)	0.64 (0.35-1.17)	0.15	
Indoor Residual Spraying							
Yes	35	3	NA	REF	NA	NA	0.17
No	1007	54	0.37	0.56 (0.17-1.94)	0.38 (0.10-1.42)	0.15	

OR* Adjusted for age only

OR** Adjusted for age, gender, care seeking, household head's education level and antimalarial use.

p-value* for age adjustment only

p-value** adjusted for age, gender, care seeking, education level and antimalarial use.

Type III p-value: P-value from a likelihood ratio test comparing a model including the variable being tested for compared to a model excluding the variable being tested for. LRT was invalid for the model including bed net quality.

3.4 Parasitological case definition of malaria

All individuals presenting for the cross-sectional survey were assessed for the presence of fever or a history of fever before a malaria blood slide was prepared. This data would assist in distinguishing between malarial fevers and non-malaria fevers. In endemic areas, asymptomatic carriage of malarial parasites is common, and apparently healthy individuals may carry parasites asymptomatically. Under such circumstances, it can be difficult to determine whether febrile episodes are attributable to malaria or other diseases. However, the likelihood that the symptoms experienced are due to malarial parasites increases with the density of parasites observed on thin blood films by microscopy. Clinical malaria can therefore be defined on the basis of fever in the presence of parasites at a density above a defined threshold value. This threshold can be determined using a logistic regression approach that estimates the probability of having fever as the dependent parameter and parasite density as the independent parameter (Smith *et al.*, 1994). Candidate threshold values are evaluated in the model to find one that gives the highest specificity and sensitivity. This value is then adopted as the threshold value that defines clinical malaria from asymptomatic malaria. The threshold also helps to distinguish if the fever symptom experienced can be attributed to malaria parasites or other possible aetiologies such as viral or bacterial agents.

Some of the questions I hoped to resolve in this study were therefore:

(i) Is malaria case definition affected by genotype?

From the foregoing, it is evident that the threshold value used as a cut-off for malaria fevers is a population average value. It is possible that the threshold may be different in particular subsets of the population, such as different genotypes or age groups. This is an

important distinction to make, as measures of protection from disease by genotypes assume a homogenous exposure. Parasite density thresholds for malaria fever were compared for significant differences across genotype and age categories.

(ii) Is the malaria attributable fraction the same across different genotypes?

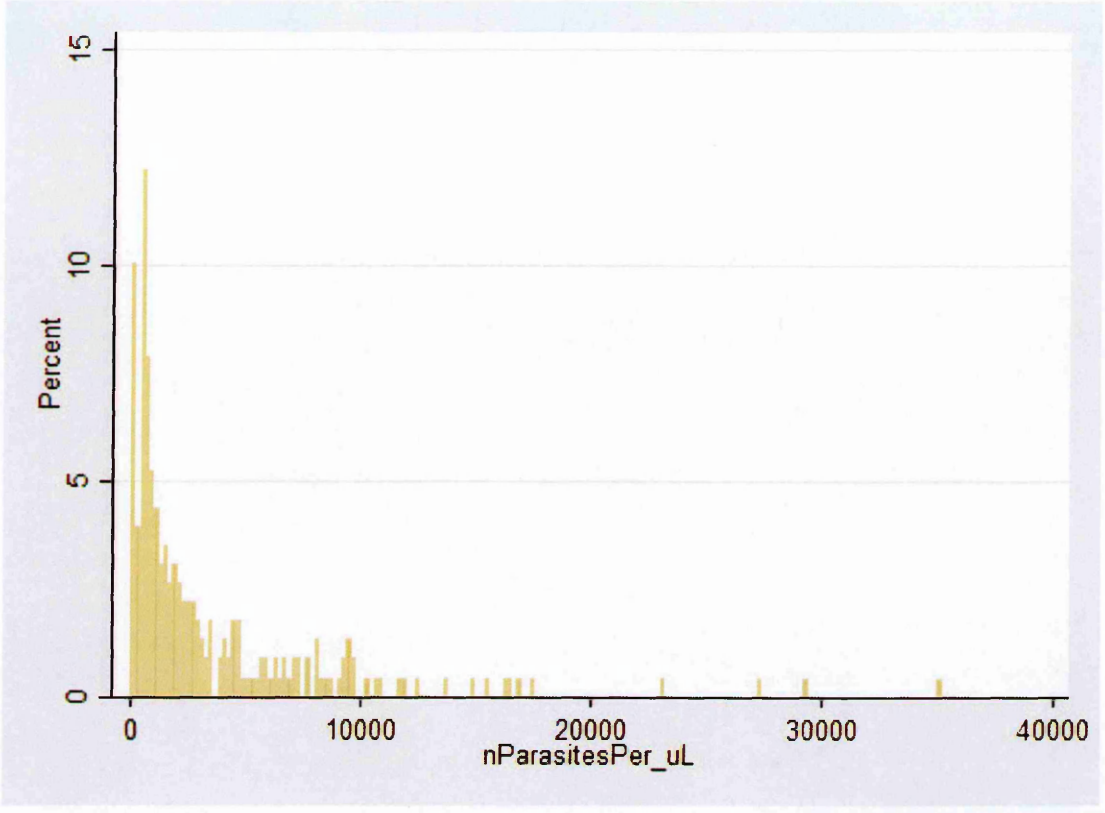
If there are significant differences in the threshold of malaria, it follows that the proportion of fevers due to malaria (the malaria attributable fraction of fever, MAF) may also be different in the various ages and genotypes. Proportions of malaria attributable fever were compared for significant differences by genotype and age categories.

To answer these questions, parasitological data was available from 1035 individuals with parasite counts ranging from 0 to 35,000 parasites / μL. The data included all the children in the age range from 0 to 12 years, of which 222 had positive slides for malaria parasites. Sections 3.4.1 onwards address these questions.

3.4.1 Distribution of slide densities

Approximately 80% of the participants had a negative malaria slide, whereas only 0.05% of the participants had a parasite density >10,000 parasites / μL. Figure 3.3 below shows a plot of the proportions of parasite density values for slide-positive individuals. Of those with non-zero values for parasite density, the majority (60%) had parasite densities less than or equal to 2,500.

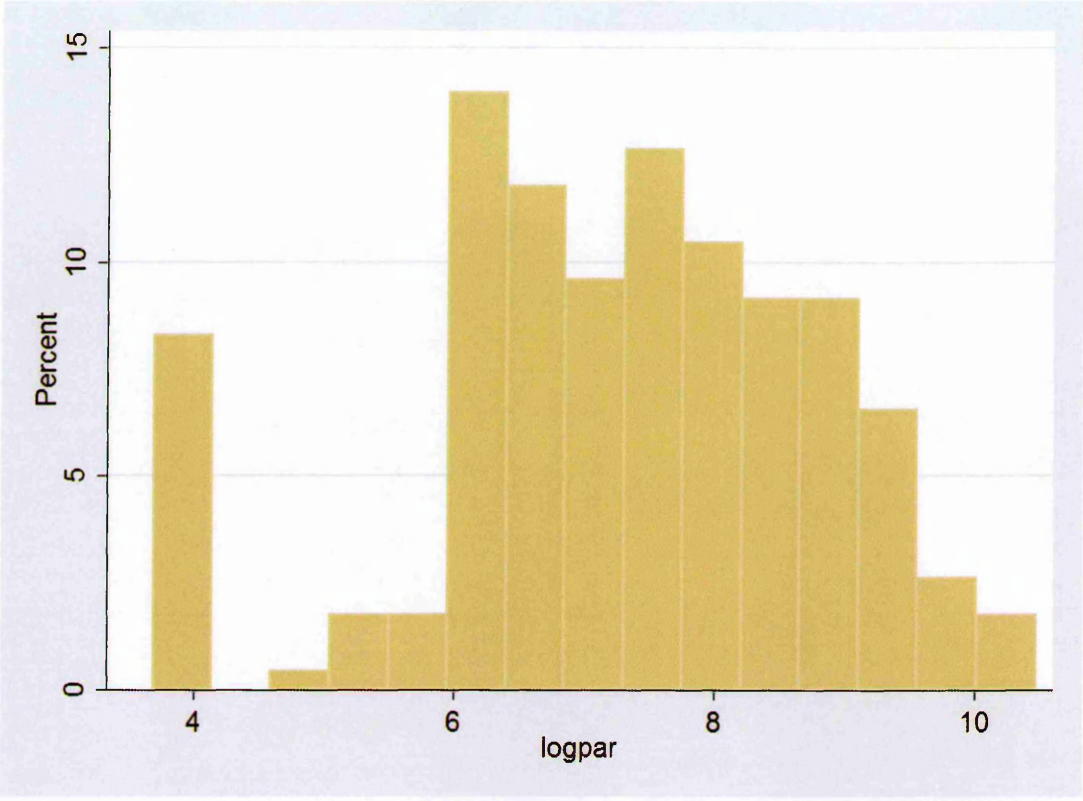
Figure 3. 3 The distribution of different parasite density values



The parasite density values are plotted in the x-axis in incremental bands of 200 parasites per µL. nParasitesPer_uL= number of parasites per µl. The y-axis represents the percentage of individuals with a particular parasite density value.

Given the skewed parasite density values, the data were transformed into natural logarithms as shown in Figure 3.4 below.

Figure 3. 4 Proportions of log transformed parasite density (logpar)



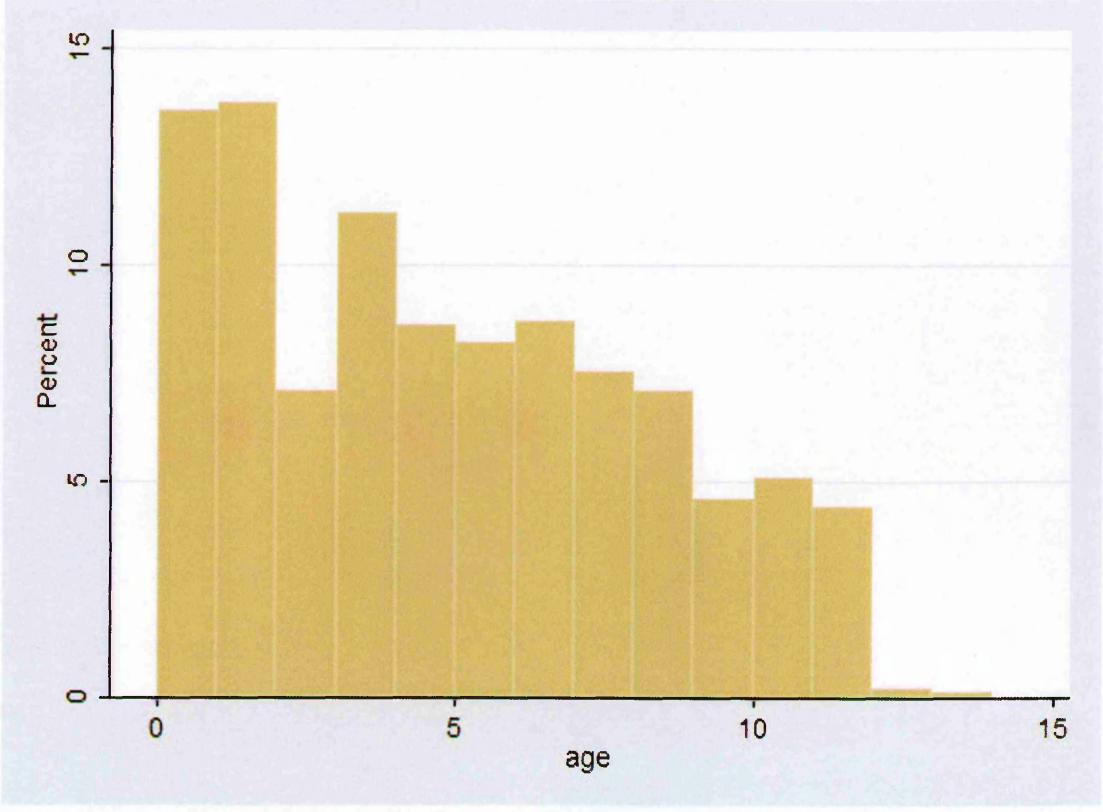
The x-axis represents different log of parasite density values whereas the y-axis is the percentage of individuals with a particular log of parasite density.

14.3% (163 individuals) had either a current fever or a fever in the last 48 hours, whereas a total of 300 individuals had a history of fever in the previous two weeks.

3.4.2 Age of malaria cross sectional survey participants

Many participants in the survey were under five years of age as shown in Figure 3.5 below.

Figure 3. 5 Age distribution of participants in cross sectional malaria survey



The x-axis denotes the age in one year age categories. The y axis shows the percentages of individuals in each age category.

3.5 Malaria attributable fraction of fever (MAF)

Fever episodes may be dichotomized into those of malarial aetiology and those that are not of malarial aetiology. The proportion of fevers due to malaria is termed the malaria

attributable fraction of fever. Determination of MAF is complicated by the fact that not all fevers in malaria parasite positive individuals are due to malaria. A plot of parasite density thresholds and their sensitivity/specificity in defining malaria fever episodes helps to determine which cut off to use. The malaria-attributable fraction at that cut-off is then determined to be the malaria attributable fraction of fever. The method employed for these calculations is described by Smith and colleagues (Smith *et al.*, 1994). Results are presented in section 3.3.3 below.

3.5.1 Attributable fractions of fever at different thresholds of parasite density

Attributable fractions are tabulated in table 3.3 below.

Table 3.3 Attributable fractions of fever at defined parasite thresholds. Sensitivity and specificity refer to the ability of the cut off definition to define a fever.

Parasite density threshold (parasites/µL)	Sensitivity	Specificity	Attributable fraction
0	100.00	0.00	0.35
500	68.43	86.52	0.73
1000	64.45	87.33	0.73
1500	56.44	88.94	0.73
2000	56.44	88.94	0.73
2500	52.43	89.74	0.73
3000	44.39	91.34	0.73
3500	44.39	91.34	0.73
4000	44.39	91.34	0.73
5000	24.26	95.30	0.73
5500	20.23	96.09	0.74
6000	20.23	96.09	0.74
6500	20.23	96.09	0.74
7000	16.19	96.88	0.74
10000	8.11	98.44	0.74
15000	8.11	98.44	0.74

From Table 3.3, it can be seen that 35% of fevers were not due to malaria parasites (AF at parasite density=0). Increase in parasite density to 500 resulted in an increase in the attributable fraction to 73%. Beyond 500 parasites / µl there was no significant difference in the attributable fraction of malaria fever, with 74% at 15,000 parasites per µl. Parasite densities were therefore investigated using a receiver operating characteristic (ROC) curve to check for a threshold that best predicts a malaria fever as shown in Figure 3.6 below. Table 3.4 shows the corresponding sensitivity and specificity at each threshold.

Figure 3.6. ROC curve showing the performance of parasite density in predicting fever episodes.

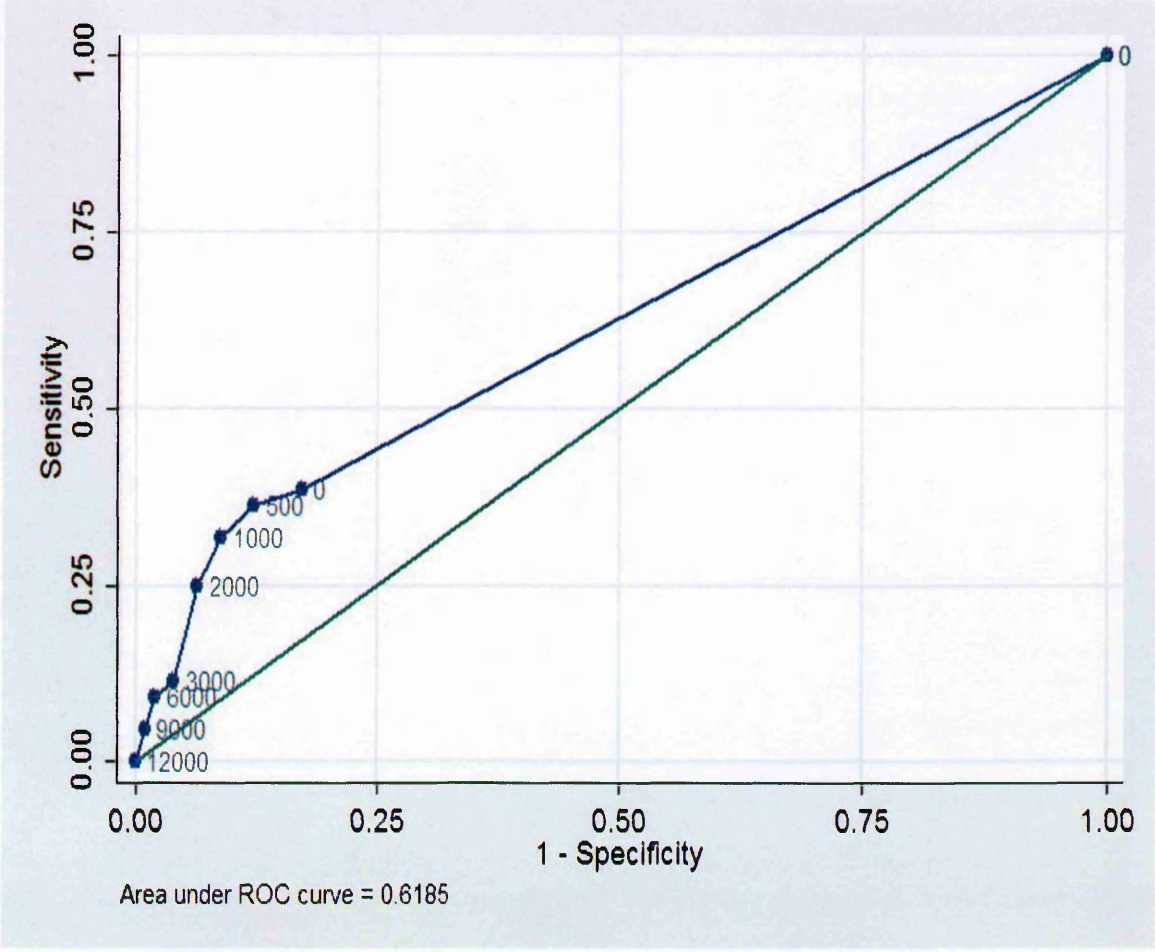


Table 3.4. Sensitivity and specificity of predicted fevers at thresholds of parasite density.

Threshold	Sensitivity %	Specificity %	Correctly classified %
>=0	100	0	4.04
>=500	38.64	82.79	81.01
>=1000	36.36	87.76	85.69
>=2000	31.82	91.2	88.81
>=3000	25	93.59	90.83
>=6000	11.36	96.08	92.66
>=9000	9.09	97.99	94.4
>=12000	4.55	99.04	95.23
>12000	0	100	95.96

The ROC curve tested the ability of parasite density to predict a fever, defined as temperature > 37.5^0C. The area under the ROC curve was 0.61 (0.53-0.68), meaning that the level of parasite density did not discriminate well individuals with malaria fever from individuals without a malaria fever. As such, these data did not show a distinct threshold that was predictive of a malaria episode. These results were consistent with the MAF calculations that did not show a difference in MAF of fever beyond a parasite density of 500. These results were unexpected since as it is commonly held in the literature that the likelihood of having malaria symptoms increases above a defined threshold of parasite density. Our results could be because there were few fever or parasitaemia episodes. Smith and colleagues (Smith *et al.*, 1994) used data collected every two months for two years from children living in an endemic area, while the current dataset was generated from a single cross sectional survey of malaria. A cross tabulation of fever and absence of fever in some indicative parasite densities is shown in Table 3.5 below. The table shows that above 2500 parasites, there were less only 13 individuals who had fever in the cohort.

This could have led to the low discriminatory ability of the ROC test to predict malaria fevers.

Table 3.5 Number of fever episodes in different parasite density cut-offs.

Parasite cut offs	Fever	No fever
>=100	19	203
< 100	27	866
>=500	17	180
< 500	29	889
>= 1500	14	108
< 1500	32	961
>=2500	13	78
< 2500	33	991
>=3500	11	58
< 3500	35	1011
>=5000	6	47
< 5000	40	1022
>=7500	4	31
< 7500	42	1038
>=10000	2	15
< 10000	44	1054

3.6 Relative burden of malaria compared with other infectious syndromes

The relative burden of malaria compared to other infectious diseases in the area was investigated by comparing the rates of malaria to the rates of other infectious disease syndromes detected in Lwak hospital. These syndromes included respiratory infections, non-malarial fevers and diarrhoea/gastroenteritis. Case definition of these syndromes was based on clinician diagnosis at presentation and the discharge diagnosis at the hospital.

Rates of uncomplicated and severe malaria were compared with respiratory rates, non-malarial fevers and diarrhoea/gastroenteritis. These are shown in Table 3.6 below.

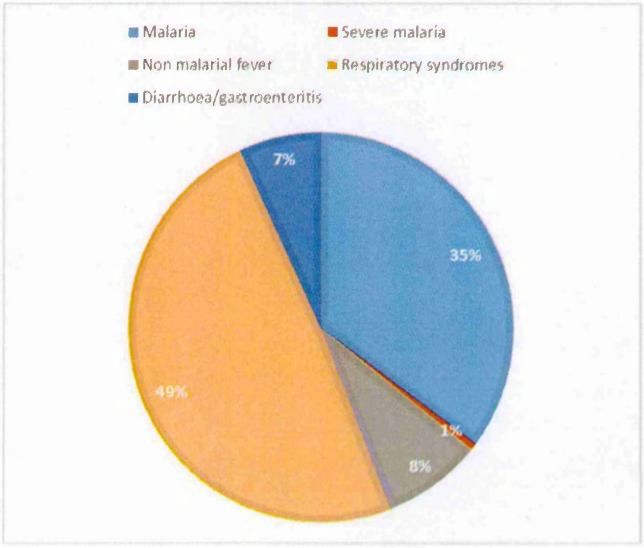
Table 3.6 Incidence of malaria compared to non-malarial illnesses diagnosed in Lwak hospital

	Number of individuals	Cyfu	Number of events	Crude Incidence (events/cyfu)* (95% CI)
Malaria	1522	5115.9	2601	0.51 (0.49-0.53)
Severe malaria	1522	5115.9	47	0.009 (0.007-0.012)
Non malarial fever	1522	5099.4	587	0.12 (0.11-0.12)
Respiratory syndromes	1524	5099.5	3643	0.71 (0.69-0.74)
Diarrhoea/Gastroenteritis	1518	5098.8	548	0.11 (0.10-0.12)

Note: Definitions of Respiratory illness, non-malaria fever and diarrhoea/gastroenteritis are explained in section 2.6.2 above.

The crude incidence rates of the infectious syndromes above were compared with crude rates of malaria in a pie-chart, as seen in Figure 3.7 below. Clinical malaria episodes accounted for about 35% of the infectious syndromes evaluated in by this study.

Figure 3.7 Malaria in comparison with infectious syndromes.



Five infectious syndromes were considered, of which uncomplicated malaria (in light blue) accounted for 35%, respiratory syndromes (orange) accounted for 49%, diarrhoea/gastroenteritis accounted for 7%, other non-malarial fever (grey) accounted for 8% and severe malaria(red) accounted for 1%.

3.6.1 Laboratory confirmation of other infectious illnesses among cohort members

3.6.1.1 Blood culture

Blood culture was performed for a subset of the cohort members to investigate the aetiology of defined clinical syndromes. Blood for culture was collected from patients who met respiratory and febrile illness case definitions. Fever was defined as an axillary temperature of $>38^0$C. Due to the expected high incidence of febrile illness in this malaria endemic area, blood for culture due to the fever criterion was only collected from the first two patients <5 years old and first two persons >5 years old presenting to the hospital with fever each day.

Respiratory illness was defined as described in section 2.6.2. Occasional individuals who did not meet these criteria had blood collected if the clinician suspected sepsis. Collection was not restricted to this genetic cohort, but to the larger PBIDS cohort of Asembo. Among the genetic cohort members, 532 individuals had blood investigations done on them. These investigations included the use of a BACTEC™ 9050 system (to detect growth of microorganisms), a Gram stain or a microbiological blood culture. A total of 1043 samples were collected during the period. Results were available for 1041 of the samples. Since individuals could be seen more than once, the total number of blood samples are more than the number of individuals contributing to the blood samples. Table 3.7 below presents results of cohort members.

Table 3.7 Blood microbiology investigation results of cohort members.

MICRO-ORGANISM	No.	%
No growth	951	91.35
Gram positive bacteria		
Streptococcus pneumoniae	5	0.48
Staphylococcus species#	12	1.15
*Gram positive bacilli***	46	4.42
Gram negative bacteria		
Haemophilus influenza	1	0.10
Klebsiella oxytoca	2	0.19
Pseudomonas oryzihabitants	1	0.10
Salmonella group H	1	0.10
Salmonella group B	12	1.15
Salmonella group D	8	0.77
Serratia rubidae	1	0.10
Escherichia coli	1	0.10
Total blood samples	1,041	100

No distinction between pathogenic and non-pathogenic Staphylococci was made for these samples

*** Gram positive bacilli were not further delineated into specific bacteria

3.6.1.2 Stool culture

Patients meeting the case definition of diarrhoea of more than 3 looser than normal stools in a day, as defined in section 2.6.2, had a stool sample collected for culture. A previous study in the study area found an incidence of diarrhoea of 0.40 episodes per year for children <5 years and 0.09 for persons ≥5 years (Feikin *et al.*, 2011). However, among the cohort members, only 36 stool samples were collected. The results were as follows:

Table 3.8 Stool culture results.

MICRO-ORGANISM	No.	%	Rotavirus result	No.	%
Gram negative bacteria					
Campylobacter coli	3	8.33	*Negative*	27	75.00
Campylobacter jejuni	3	8.33	*Positive*	8	22.22
Salmonella Group B	1	2.78	*Not tested*	1	2.78
Salmonella species	1	2.78			
Shigella flexneri	4	11.11			
Shigella sonnei	2	5.56			
Vibrio cholera	2	5.56			
Gram positive bacteria	0	0			
No pathogen detected	22	61.11			
Total	**36**	**100**	**Total**	**36**	**100**

Percentages represent the proportion of samples that grew the microorganism out of the total number of samples tested. The total is greater than 100% because two participants had more than one microorganism isolated. One had *Shigella Sonnei* and *Campylobacter Coli* whereas another individual had *Salmonella species* and *Campylobacter Coli.*

No pathogen detected consists of normal human gut flora.

3.6.1.3 Nasopharyngeal and oropharyngeal swab tests

Upper and lower respiratory tract infections may result from several aetiologies, including *Influenza* viruses. Among the cohort members, it was possible to capture the data for nasopharyngeal oropharyngeal swabs tested for *Influenza A* and *B*. The results are presented in Table 3.9 below. Of 608 samples collected, there were 29 positive for *Influenza A* (4.8%), whereas 14 (2.5%) were positive for *Influenza B*.

Table 3.9 Results for Nasopharyngeal oropharyngeal swabs.

Flu A Result		**No.**	**%**	**Flu B Result**	**No.**	**%**
Not determined		2	0.33	*Negative*	593	97.5
Negative		565	92.93	*Positive*	15	2.5
Positive		41	6.74			
Positive Sub typed		29	4.77			
	H1	3	0.49			
	H3	8	1.32			
	NH1	14	2.30			
	SUB INC*	4	0.66			
Total		**608**			**608**	**100**

H1=Haemagglutinin 1
H3=Haemagglutinin 3
NH1=Neuraminidase Haemagglutinin 1
SUB INC= subtype not confirmed

Sections 3.6.1.1 through 3.6.1.3 above show that individually each aetiological agent contributed only a small portion to the disease incidence recorded. However, together, the various aetiological agents contributed to the infectious disease syndromes such as respiratory illness and diarrhoea in this cohort. Since not all individuals presenting with illness were sampled, incidence due to a particular aetiological agent could not be estimated without adjusting for the individuals that were not sampled. Going forward, investigation of these other syndromes were not investigated as a blood sample was not available systematically from these cohort members. The information on which aetiological agents cause the infectious syndromes was nevertheless useful to show the relative importance of the various aetiological agents in causing illness in the study area compared to *P. falciparum* malaria.

3.7 Discussion

As expected, the malaria incidence within the cohort study area increased by age from birth up to 6 years, after which it declined, an observation that almost certainly reflects the acquisition of malaria-specific immunity. Younger children are expected to have low immunity, and therefore malaria affects them more. It appears that beyond 5 years the immunity is well developed and consequently less malaria is experienced by the children.

Since this was a hospital-based study, the rates reported are likely to be representative of only those individuals who sought care at the hospital. One of the major factors which may affect the decision to seek care at a hospital is the distance from the patient's home to the hospital. A previous study in the area had found that rates of fever in the home were approximately 10 times higher than in those seen in the hospital. This may explain the low rates seen in this study of about 0.5 episodes / cyfu, compared to the rates recently reported during the RTS,S malaria vaccine study that took place approximately within the same period (2009-2013) which reported rates approximately 10 times higher, of 4-5 episodes / cyfu.

Recent declines in malaria have been reported globally. The rates of malaria in the study area over the period appeared to be constant compared to the reference year, 2008. However, 2 years stand out as having different rates, an increase as seen in 2010 and a decline reported in 2013. A nested cross sectional cohort study was conducted during the study for two reasons - to derive a parasitological definition and to investigate risk factors for having malaria parasite. While it was possible to define a number of risk factors, such as age and effect of care seeking on malaria, a clear cut-off defining fever was not evident. This could have been due to the low number of measured fever events during the cross

sectional survey. In addition, a review of fever and malaria has shown that fevers due to malaria are declining, especially in East Africa (D'Acremont *et al.*, 2010). This has implications for presumptive treatment of fevers in areas classified as endemic, where treatment should be guided by evidence of a malaria parasite positive slide.

3.5 Summary

1. Malaria is among the major of infectious illnesses reported in Lwak hospital in this study area, representing a significant health burden.

2. Major infectious disease syndromes in the study area include respiratory tract infections, diarrhoea and non-malarial fevers. Data was only available in some children for definitive diagnosis of aetiological agents in these syndromes, and therefore conclusive analysis of incidence by aetiological agent was not done.

3. Risk factors for malaria illness has been investigated through a longitudinal approach, showing the that time-changing variables such as season, age, calendar year, season and distance to hospital affect rates of malaria reported. Investigation through a cross sectional survey showed a significant effect for participants age and whether the participant sought care, whereas condition of bed net, house type, gender and distance from hospital did not affect the odds of having malaria.

4. The parasitological case definition of malaria was investigated in this population using cross sectional malaria survey of malaria symptoms such as fever compared with parasite density of malaria. From the data analysis, a clear threshold was not seen.

5. The rates calculated in this study were significantly influenced by distance to the health facility. The odds of attending the surveillance hospital was found to change significantly for every 2km.

CHAPTER 4: Red blood cell genes and malaria incidence in Asembo

4.1 Background

The classical genes associated with protection from malaria are related to the red blood cell. These polymorphisms include the sickle cell trait, α-thalassaemia, the ABO blood groups system, G6PD deficiency and polymorphisms in the CR1 Knops blood group system, particularly at the Swain Langley and McCoy loci. However, from a survey of the literature (see Chapter 1), there is no consensus on the effects of some of these genes, such as G6PD deficiency and CR1 polymorphisms. While several studies have focused on the effects of these genes on severe malaria, their effect on uncomplicated malaria has only been reported in relatively few studies. In this chapter, variants of these common red blood cell gene polymorphisms are investigated for their effect on malaria incidence in the children of Asembo. Incidence data is further stratified by gene variant and age to give indications of the rate of development of naturally acquired immunity in the various variants. The investigations in this chapter will therefore aim to answer the following questions:

1. Do polymorphisms in the classical red blood cell genes affect the incidence of uncomplicated clinical malaria?
2. Do the age-specific rates provide evidence for a difference in the rate of acquisition of naturally acquired immunity in the different genotypes?
3. Do red blood cell polymorphisms affect the parasitological case definition of malaria?

4.2 Objectives

Specific objectives were as follows:

1. To investigate protection conferred against malaria incidence by red blood cell genes in Asembo.
2. To investigate the age specific malaria incidence patterns of the red blood cell polymorphisms in order to infer associations with naturally acquired immunity.
3. To investigate the role of common red blood cell genes in predicting malaria fevers.

4.3 Summary of methods

Methods for surveillance, case definitions for malaria and the statistical methods have been described in detail in Chapter 2, summarised briefly as follows. Passive hospital based surveillance in Lwak hospital was employed to follow up enrolled children for incident malaria episodes between 2008 and 2013. Crude and adjusted incidence rates were calculated by Poisson regression. Crude incidence rates were calculated by dividing the diseases episodes by the follow-up time. Adjusted rates were calculated by adding to the Poisson model other variables found to significantly affect the incidence of malaria on univariate analysis. These variables were season, year, distance to Lwak hospital, age, gender, mother's education and socio-economic status. An additional adjustment was using the sandwich estimator to account for repeated observations on the same individual. This adjustment reduces the p-values and inflates the corresponding confidence intervals. Age and genotype specific incidence rates and incidence rate ratios were calculated by restricting the model to individuals of particular ages and genotypes respectively.

Likelihood ratio test for effect of age on genotype specific rates

To further investigate if the interaction between genotype and age is important in predicting malaria episodes, I conducted a likelihood ratio test to check for a statistically significant interaction between genotype and age. I reasoned that if the interaction between genotype and age is important, then it should significantly improve a Poisson model which predicts malaria from its covariates. I therefore compared two models for each red blood cell genetic polymorphism:

Model (i) a Poisson regression model with a red blood cell genotype, age category and the covariates season, year, distance to Lwak hospital, age, gender, mother's education and socio-economic status.

Model (ii) Poisson regression model similar to model (i) but in addition having an interaction term for age category and sickle genotype among the independent variables. A likelihood ratio test comparing the likelihood estimates from the two models was applied, and significance for differences between the estimates was evaluated at the $P<0.05$ level.

4.4 Results

4.4.1 Sickle

4.4.1.1 Effect of sickle on incidence of uncomplicated malaria

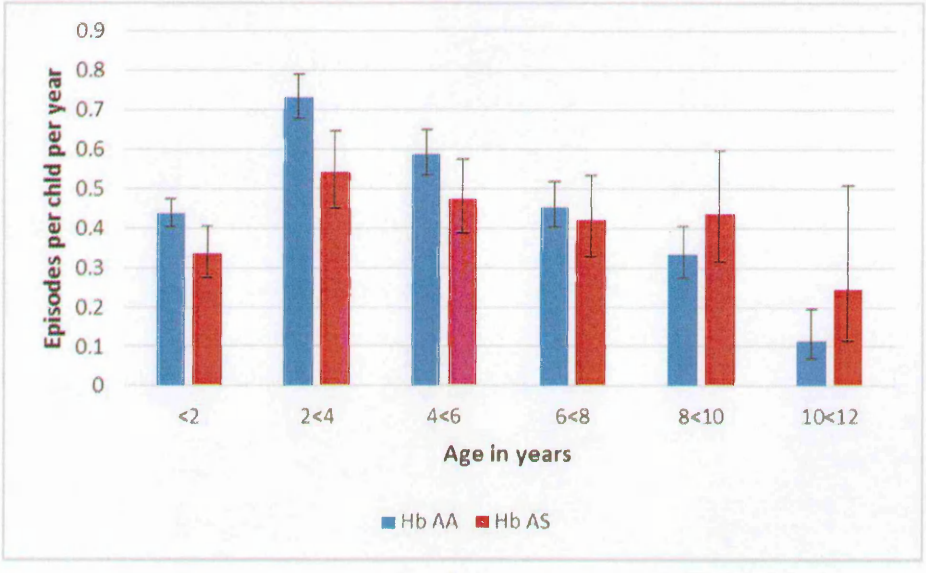
Out of 1444 participants, 1132 (78%) had HbAA, 289 (20%) had HbAS and 23 (1.6%) had HbSS. The overall incidence rate ratio for malaria in children with HbAS compared to HbAA among children 0-12 years old was 0.83 (0.75-0.92; $p<0.001$). This reduction

was significant both on crude analysis and on adjustment for season, year, distance (km) to Lwak hospital, age, gender, mother's education and socio-economic status. However, upon additional adjustment of p-values and confidence intervals using the sandwich estimator, the protective effect was not evident, with IRR 0.85 (0.69-1.04; p=0.12).

4.4.1.2 Age-specific incidence of malaria by sickle cell genotypes

Differences in the malaria age-incidence curves were investigated by comparing the age-specific incidence rates in normal children, HbAA, compared children with HbAS (Table 4.1). The incidence of malaria by genotype and age is presented in Figure 4.1 below. The rates in the stratified age categories were not significantly different compared to the overall effect in all the age categories combined.

Figure 4.1 Rates of uncomplicated malaria in HbAA compared to HbAS individuals stratified by age categories.



The y-axis represents the episodes of malaria per child per year of follow-up, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children who are normal at the sickle cell locus (HbAA), while the red bars represent children who are carriers for the sickle cell trait, HbAS.

Malaria incidence for HbAA formed a clear peak incidence at the 2<4 year category, compared to the preceding 0<2 years, and the next age category, the 4<6 year category. By contrast, the HbAS peak incidence at 2<4 years appears to be significantly different from only the preceding age category's (0<2) and but not the other age categories above 4 years. Malaria incidence also declines more sharply for older age groups of HbAA compared to HbAS. Interestingly, there is a higher incidence of malaria in AS children compared to AA children in the 8-10 and >10 year age categories, although the difference did not reach statistical significance. These two phenomena, the peak shift and apparent shift in malaria protection in the older age groups suggest that there may be a difference in the acquisition rate of naturally acquired immunity in the HbAA compared to the HbAS children.

A likelihood ratio test comparing the age interaction model and the non-age interaction model gave a p-value of 0.045, indicating marginal significance at the <0.05 threshold. This suggests that age interacts with genotype to predict rates of malaria. Since the effect of age may be heterogeneous across the age categories, the rates of malaria were presented in each age category in Table 4.1 along with the number of malaria episodes, number of individuals per genotype in each age category, the child years of follow up and incidence rate ratio adjusted for the covariates.

Table 4.1 Age-specific incidence of malaria and incidence rate ratios by sickle genotype.

Age range (Years)	Genotypes	No. of participants (a)	No. of malaria episodes	cyfu	Crude malaria incidence (b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
All	AA	1130	1966	3814.82	0.52	1		1	
	AS	286	442	1032.50	0.43	0.83 (0.75-0.92)	<0.001	0.85(0.69-1.04)	0.12
0-2	AA	932	582	1300.00	0.44	1		1	
	AS	226	106	317.28	0.33	0.76 (0.62-0.94)	0.01	0.78 (0.56-1.09)	0.15
2-4	AA	516	626	854.40	0.73	1		1	
	AS	137	123	227.05	0.54	0.74 (0.61-0.90)	<0.01	0.75 (0.56-1.01)	0.06
4-6	AA	471	407	691.00	0.59	1		1	
	AS	134	102	215.00	0.47	0.81 (0.65-1.00)	0.05	0.89 (0.61-1.30)	0.55
6-8	AA	325	239	524.67	0.46	1		1	
	AS	100	66	157.06	0.42	0.92 (0.70-1.21)	0.56	0.92 (0.61-1.38)	0.68
8-10	AA	200	98	293.35	0.33	1		1	
	AS	59	38	87.35	0.44	1.30 (0.90-1.89)	0.17	1.20 (0.68-2.10)	0.53
>=10	AA	101	14	119.95	0.12	1		1	
	AS	28	7	28.76	0.24	2.09 (0.84-5.17)	0.11	2.15 (0.48-9.65)	0.32

4.4.2 Alpha thalassaemia

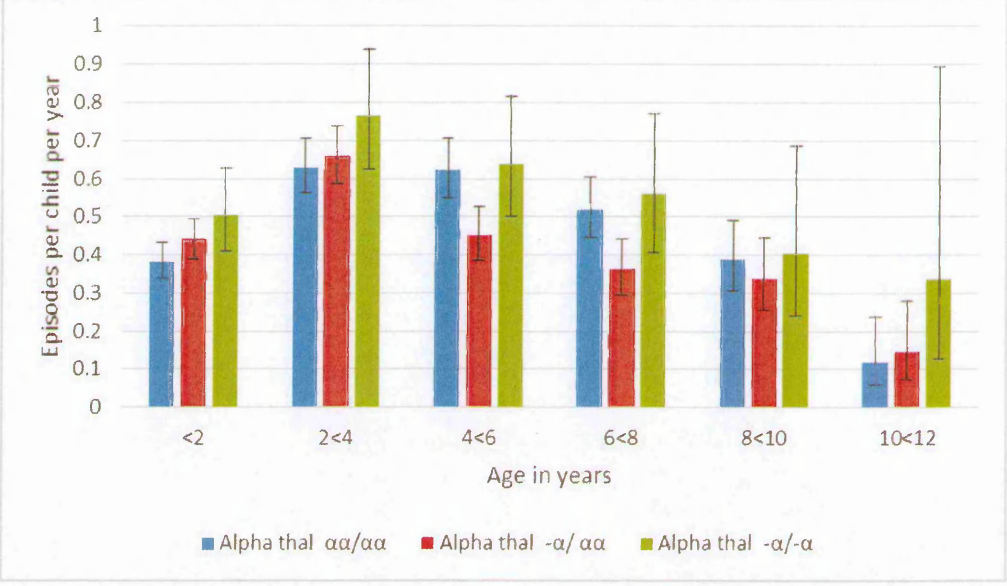
4.4.2.1 Effect of α–thalassaemia on incidence of malaria

Genotyping for α^{+}thalassaemia was completed for 1299 individuals. There were 616 (47.4%) normal (αα/αα) individuals, 547 (42.1%) heterozygotes (-α/αα) and 136 (10.5%) homozygotes (-α/-α). The frequency of the $\alpha^{-3.7}$ allele was 0.32. The crude malaria incidence rate ratio in α^{+}thalassaemia homozygotes compared to normal individuals was significantly higher, with IRR 1.20 (1.05-1.36; $p<0.01$). However, upon adjustment for covariates and the sandwich estimator, the significant effect was lost, with IRR 1.01 (0.76-1.35; $p=0.92$). Heterozygous individuals showed no differences in either the crude analysis (IRR 0.94 CI 0.86-1.03, $p=0.17$) or the adjusted analysis (IRR 1.02, CI 0.85-1.23, $p=0.83$).

4.4.2.2 Age-specific incidence of malaria by α-thalassaemia genotype

The age-stratified effects of α-thalassaemia are shown in Figure 4.2 and Table 4.2 below. The adjusted analyses did not show an effect of α-thalassaemia specific to any age category, as seen in Table 4.2.

Figure 4.2 Rates of uncomplicated malaria in α -thalassaemia genotypes stratified by age categories.



The y-axis represents the episodes of malaria per cyfu, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children who normal at the alpha thalassaemia locus (αα/αα), orange bars represent children who are heterozygous (-α/αα) whereas the grey bars represent children who are homozygous (-α/-α).

A likelihood ratio test comparing the α-thalassaemia genotype and age interaction model with the non-age interaction model gave a p-value of 0.0083. Age therefore interacts with thalassaemia genotypes to predict malaria incidence rates. The age-specific incidence of malaria by α-thalassaemia genotype is presented in Table 4.2 below.

Table 4.2 Age-specific malaria incidence rates stratified by α^+thalassaemia genotypes.

Age range (Years)	Genotypes	No. of participants (a)	No. of malaria episodes	cyfu	Crude malaria incidence (b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
All	αα/αα	616	1061	2144.10	0.49	1		1	
	-α/αα	547	902	1940.71	0.46	0.94 (0.86-1.03)	0.17	1.02 (0.85-1.23)	0.83
	-α/-α	136	297	501.87	0.59	1.20 (1.05-1.36)	<0.01	1.01 (0.76-1.35)	0.92
0-2	αα/αα	495	270	704.73	0.38	1		1	
	-α/αα	447	291	661.75	0.44	1.15 (0.97-1.35)	0.10	1.27 (0.93-1.73)	0.13
	-α/-α	110	84	166.07	0.51	1.32 (1.03-1.69)	0.03	1.16 (0.77-1.74)	0.48
2-4	αα/αα	285	296	470.71	0.63	1		1	
	-α/αα	268	294	445.65	0.66	1.05 (0.89-1.23)	0.56	1.14 (0.89-1.48)	0.3
	-α/-α	72	92	120.02	0.77	1.22 (0.96-1.54)	0.10	1.04 (0.68-1.61)	0.85
4-6	αα/αα	273	253	405.76	0.62	1		1	
	-α/αα	242	162	359.32	0.45	0.72 (0.59-0.88)	<0.01	0.78 (0.57-1.05)	0.1
	-α/-α	66	65	101.46	0.64	1.03 (0.78-1.35)	0.85	0.76 (0.44-1.31)	0.33
6-8	αα/αα	195	164	315.12	0.52	1		1	
	-α/αα	168	96	264.37	0.36	0.70 (0.54-0.90)	<0.01	0.72 (0.49-1.06)	0.1
	-α/-α	46	38	67.90	0.56	1.08 (0.76-1.53)	0.69	1.00 (0.62-1.60)	1.00
8-10	αα/αα	121	70	180.28	0.39	1		1	
	-α/αα	100	50	147.77	0.34	0.87(0.61-1.25)	0.46	1.08 (0.61-1.93)	0.78
	-α/-α	25	14	34.49	0.41	1.05 (0.59-1.86)	0.88	1.40 (0.55-3.55)	0.49
>=10	αα/αα	63	8	67.50	0.12	1		1	
	-α/αα	51	9	61.86	0.15	1.23 (0.47-3.18)	0.67	2.20 (0.61-7.90)	0.23
	-α/-α	10	4	11.93	0.34	2.83 (0.85-9.39)	0.09	5.71 (1.41-23.17)	0.01

4.4.3 ABO blood groups

4.4.3.1 The effect of ABO blood groups on incidence of uncomplicated malaria

Compared to blood group O, the incidence of malaria in children with blood groups A, B and AB was higher on the crude analyses. Blood group A gave marginal evidence of increase in malaria incidence with an IRR of 1.10 (1.00-1.22; p=0.06). Blood group B was associated with an IRR of 1.27 (1.15-1.40; p<0.0001), while the IRR in blood group AB children was 1.47 (1.22-1.77; p<0.0001). However, none of these blood groups maintained significance for effect after adjustment for covariates and robust analysis using the sandwich estimator.

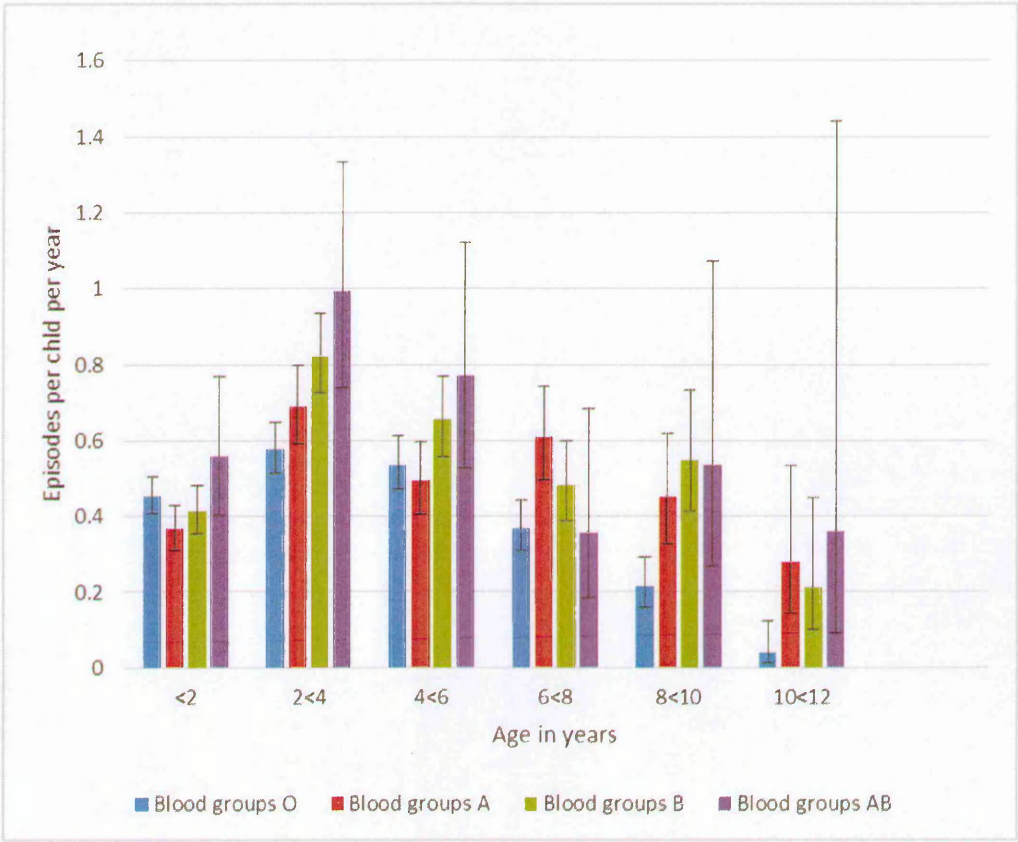
Blood group O compared to the non-O blood groups had a reduced malaria incidence, with IRR 0.88 (0.80-0.96; p=0.003) when adjusted for covariates. However, when an additional adjustment was made using the sandwich estimator, the protective effect was not evident (0.74-1.04; p=0.13).

4.4.3.2 Age specific incidence of malaria by ABO blood groups genotypes

ABO blood groups affect the age-specific incidence of malaria. Above the age of 6 years there was a significant difference between blood group O and blood group A in all age categories considered in the crude analysis. However, in the adjusted analyses, there was a protective effect of blood group O only in children above 8 years of age when compared to particular ABO genotypes. The IRR in blood group B compared to O was 2.30 (95% CI 1.2-4.4; P=0.01). Among children ≥10, the blood group A individuals had an IRR of 5.65 (95% CI 1.42-22.5; P=0.01), whereas blood group B had an IRR of 5.02 (95% CI 1.34-18.87; P=0.02). However, these findings should be interpreted cautiously given the low number of individuals in these age categories. The effect of age on malaria incidence

is shown in Figure 4.3 and Table 4.3 below. Note that Figure 4.3 plots the crude age effects unadjusted for covariates.

Figure 4.3. Rates of uncomplicated malaria in ABO blood groups stratified by age categories.



The y-axis represents the episodes of malaria per cyfu, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children with blood group O, red bars represent children with blood group A, green bars represent blood group B, whereas the purple bars represent children with blood group AB.

A likelihood ratio test comparing the ABO genotype and age interaction model with the non-age interaction model gave a p-value of <0.0001. Age therefore interacts with ABO blood groups to predict malaria incidence rates. Age specific incidence rates of malaria by ABO blood groups are presented in Table 4.3 below.

Table 4.3. Age specific malaria incidence stratified by ABO blood groups.

Age range	Genotypes	No. of participants(a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
(Years)									
All	O	641	1014	2250.72	0.45	1		1	
	A	355	573	1153.06	0.50	1.10 (1.00-1.22)	0.06	1.10 (0.90-1.34)	0.35
	B	342	127	191.36	0.66	1.27 (1.15-1.40)	<0.0001	1.19 (0.97-1.45)	0.09
	AB	59	695	1213.60	0.57	1.47 (1.22-1.77)	<0.0001	1.06 (0.65-1.73)	0.80
0-2	O	510	331	728.45	0.45	1		1	
	A	298	152	414.61	0.37	0.81 (0.67-0.98)	0.03	0.86 (0.62-1.19)	0.37
	B	278	167	403.43	0.41	0.91 (0.76-1.10)	0.33	0.88 (0.64-1.20)	0.41
	AB	50	37	66.30	0.56	1.23 (0.87-1.73)	0.24	0.84 (0.44-1.60)	0.59
2-4	O	299	284	491.62	0.58	1		1	
	A	153	174	252.31	0.69	1.19 (0.99-1.44)	0.07	1.20 (0.88-1.62)	0.25
	B	174	241	292.26	0.82	1.43 (1.20-1.69)	<0.0001	1.31 (0.99-1.73)	0.06
	AB	28	44	44.29	0.99	1.72 (1.25-2.36)	<0.001	1.18 (0.62-2.26)	0.62
4-6	O	293	231	429.81	0.54	1		1	
	A	136	104	211.15	0.49	0.92 (0.73-1.16)	0.46	0.98 (0.70-1.38)	0.91
	B	154	152	232.05	0.66	1.22 (0.99-1.50)	0.06	1.20 (0.85-1.69)	0.3
	AB	24	27	35.08	0.77	1.43 (0.96-2.13)	0.08	1.27 (0.65-2.49)	0.49
6-8	O	203	123	332.03	0.37	1		1	
	A	104	95	156.34	0.61	1.64 (1.26-2.14)	<0.001	1.40 (0.95-2.08)	0.09
	B	105	81	167.71	0.48	1.30 (0.98-1.73)	0.06	1.29 (0.87-1.93)	0.21
	AB	16	9	25.24	0.36	0.96 (0.49-1.89)	0.91	0.78 (0.29-2.07)	0.61

Age range	Genotypes	No. of participants(a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
8-10	O	131	42	193.12	0.22	1		1	
	A	57	39	86.25	0.45	2.08 (1.34-3.22)	<0.001	1.97 (1.04-3.74)	0.04
	B	63	47	85.46	0.55	2.53 (1.67-3.83)	<0.0001	2.30 (1.20-4.40)	0.01
	AB	9	8	14.90	0.54	2.47 (1.16-5.26)	0.02	0.94 (0.35-2.53)	0.91
>=10	O	67	3	75.68	0.04	1		1	
	A	28	7	32.70	0.21	7.01 (1.90-25.88)	<0.01	5.65(1.42-22.55)	0.01
	B	27	11	39.40	0.28	5.40 (1.40-20.88)	0.01	5.02 (1.34-18.87)	0.02
	AB	6	2	5.54	0.36	9.10(1.40-20.88)	0.01	NA	NA

4.4.4 G6PD deficiency

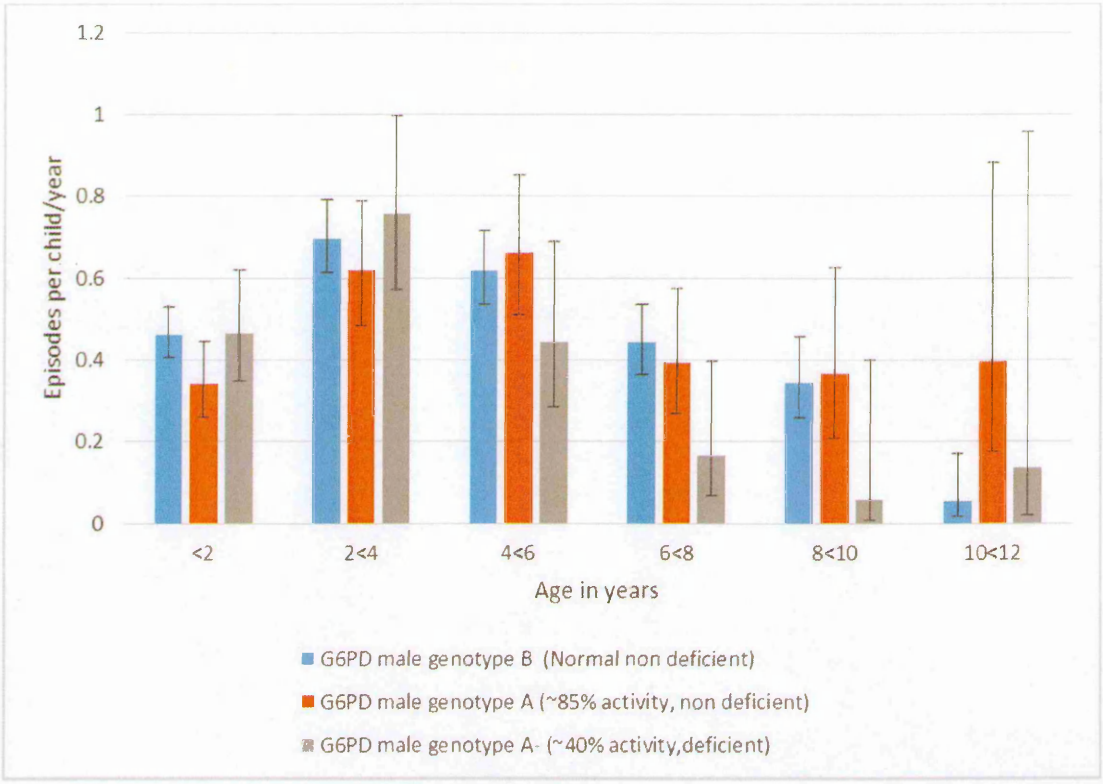
4.4.4.1 Effect of G6PD deficiency on incidence of malaria in males

There was no effect of G6PD genotypes on malaria incidence, since neither the crude or adjusted analysis gave an effect. In the adjusted analysis, G6PD A individuals had an IRR of 1.02 (0.79-1.32; p=0.89) while G6PD A- individuals had an IRR of 0.86 (0.57-1.30; p=0.46).

4.4.4.2 Age specific incidence of malaria by G6PD genotypes in males

Age stratified analyses did not show any significant effect of G6PD genotypes on malaria incidence in any age category considered in the adjusted analyses. These effects are shown in Figure 4.4 and Table 4.4 below. The crude analyses had suggested a marginal effect of G6PD A (376) in the 0-2 year old children. Crude analyses had also indicated an effect for G6PD A (376) in children >10 years, but the low number of children in this age category indicate that this result should be cautiously interpreted.

Figure 4.4. Rates of uncomplicated malaria in G6PD genotypes of males stratified by age categories.



The y-axis represents the episodes of malaria per cyfu, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children with G6PD genotype B, orange bars represent children with G6PD genotype A, while grey bars represent children with G6PD genotype A-.

The LRT comparing the interaction between G6PD male genotypes and age with the non-genotype-age interaction model resulted in a p-value of 0.03. This suggests that age interacts with G6PD genotype in males to predict malaria incidence rates.

Age specific incidence rates of malaria by G6PD genotypes are presented in Table 4.4 below.

Table 4.4. Age specific malaria incidence rates stratified by G6PD deficiency genotypes in males

Age range	Genotypes	No. of participants(a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
All	B	435	783	1517.05	0.52	1		1	
	A	145	223	470.93	0.47	0.92 (0.79-1.06)	0.26	1.02 (0.79-1.32)	0.89
	A-	86	124	267.67	0.46	0.90 (0.74-1.08)	0.26	0.86 (0.57-1.30)	0.46
0-2	B	330	214	462.64	0.46	1		1	
	A	115	54	158.39	0.34	0.74 (0.55-0.99)	<0.05	0.78 (0.52-1.18)	0.24
	A-	73	47	101.09	0.46	1.01 (0.73-1.38)	0.97	0.77 (0.40-1.46)	0.42
2-4	B	207	234	336.03	0.70	1		1	
	A	63	64	103.49	0.62	0.89 (0.67-1.17)	0.28	1.05 (0.72-1.55)	0.79
	A-	40	50	66.13	0.76	1.09 (0.80-1.47)	0.53	1.09 (0.65-1.81)	0.75
4-6	B	202	182	294.38	0.62	1		1	
	A	61	59	89.44	0.66	1.07 (0.80-1.43)	0.67	1.03 (0.65-1.63)	0.91
	A-	35	20	45.07	0.44	0.72 (0.45-1.14)	0.16	0.79 (0.45-1.40)	0.42
6-8	B	146	103	232.95	0.44	1		1	
	A	45	27	68.69	0.39	0.89 (0.58-1.36)	0.59	0.91 (0.50-1.66)	0.76
	A-	21	5	30.23	0.17	0.37 (0.15-0.92)	0.03	0.65 (0.29-1.46)	0.3
8-10	B	93	47	136.63	0.34	1		1	
	A	25	13	35.76	0.36	1.06 (0.57-1.95)	0.86	1.44 (0.61-3.40)	0.40
	A-	10	1	17.75	0.06	0.16 (0.02-1.19)	0.07	0.18 (0.03-1.14)	0.07
>=10	B	48	3	54.41	0.06	1		1	–
	A	11	6	15.16	0.40	7.18(1.80-28.70)	<0.01	NA	NA
	A-	6	1	7.40	0.14	2.45 (0.25-23.55)	0.44	3.34(0.69-16.01)	0.13

4.4.4.3 Effect of G6PD deficiency on incidence of malaria in females

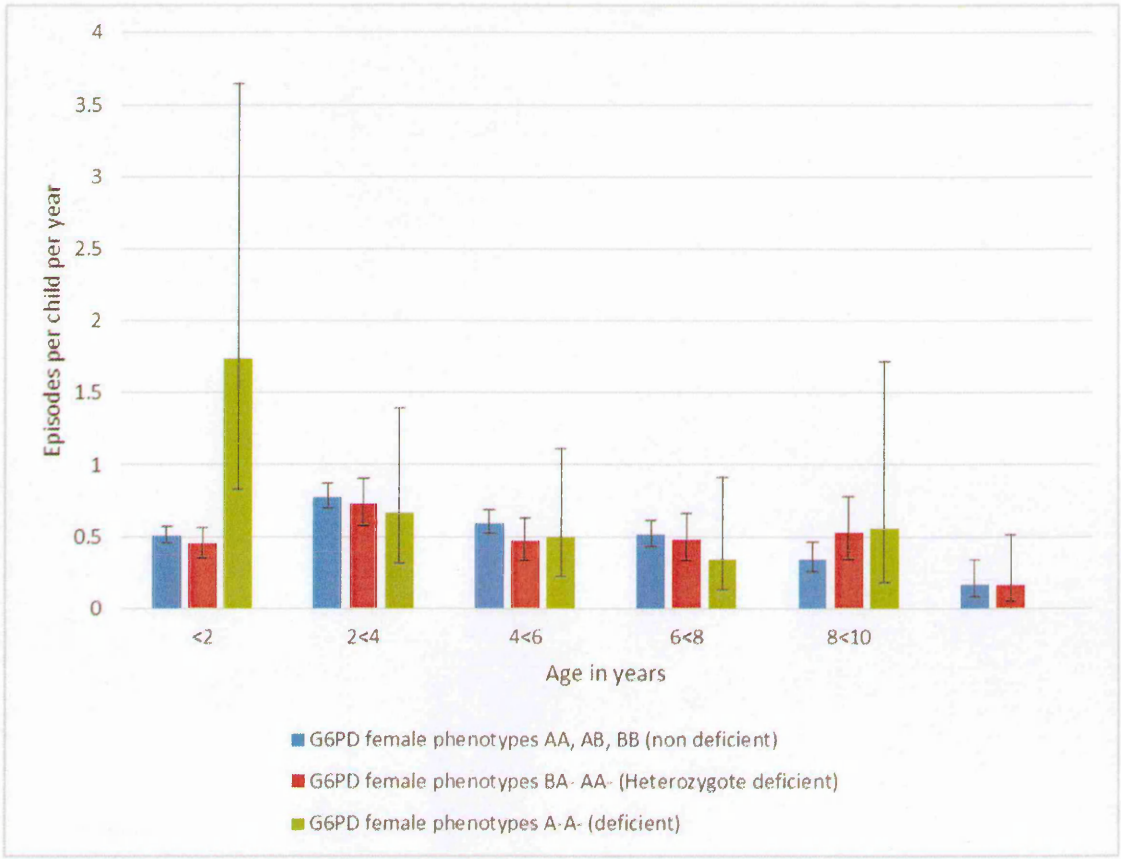
When all age categories were analysed in combination, no effect of G6PD deficiency was seen in females. Heterozygous deficient individuals AA- and BA- were grouped together. Compared with the non-deficient group (BB, BA and AA), the IRR for malaria in the heterozygous deficient group was 0.94 (0.70-1.25; p=0.66). When the deficient group (A-A-) was compared to the non-deficient group, the IRR was 0.82 (0.44-1.53; p=0.54).

4.4.4.4 Age specific incidence of malaria by G6PD genotypes in females

When the results were analysed by age categories, a significant increase in malaria incidence was seen among the youngest A-A- (deficient) children 0-2 years old. The IRR was 5.22 (1.95-13.97; p<0.01). However, this finding may not be generalizable as only 3 participants had A-A- in this age category, and therefore any malaria incident easily inflates the IRR compared to the AA wild type. In the 8-10 years category, A-A- individuals also had a higher IRR 2.76 (95% CI 1.19-6.64) compared to non-deficient (BB, BA and AA) individuals. Only 5 individuals were in the 8-10 year category and therefore the finding should be interpreted with caution. The effect of G6PD female genotypes on malaria are shown in Table 4.5 and Figure 4.5 below.

Overall, comparing the trends of malaria incidence according to G6PD deficient genotypes among males and the females (Table 4.4 for males and Table 4.5 for females), it appears the deficient genotype increases malaria incidence in some age groups among females but not in males. Considering the low numbers of individuals as discussed above, this finding may not be generalizable and would need to be confirmed in adequately powered studies.

Figure 4.5. Rates of uncomplicated malaria by G6PD genotype in females stratified by age categories.



The y-axis represents the episodes of malaria per cyfu, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children with G6PD genotype AA, AB, BB (termed non-deficient). Red bars represent heterozygote deficient individuals with genotypes BA- and AA-. Green bars represent children with G6PD genotype A-A-(deficient).

The LRT comparing the interaction between G6PD genotypes and age with the non-genotype-age interaction model had a p-value of 0.07. There was therefore no evidence that age interacts with G6PD phenotypes in females to predict malaria incidence rates. Age specific incidence of malaria stratified by G6PD genotypes are presented in Table 4.5 below.

Table 4.5. Age-specific malaria incidence rates stratified by G6PD deficiency genotypes in females.

Age range	Genotypes	No. of participants (a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
(Years)									
All	BB, BA, AA (non-deficient)	524	981	1734.92	0.57	1		1	
	AA- ,BA- (Heterozygous deficient)	148	256	498.55	0.51	0.91 (0.79-1.04)	0.17	0.94 (0.70-1.25)	0.66
	A-A- (deficient)	8	27	46.03	0.59	1.04(0.71-1.52)	0.85	0.82(0.44-1.53)	0.54
0-2	BB, BA, AA (non-deficient)	415	290	569.11	0.51	1		1	
	AA- BA- (Heterozygous deficient)	115	74	164.78	0.45	0.88 (0.68-1.14)	0.33	1.00 (0.64-1.56)	0.99
	A-A- (deficient)	3	7	4.02	1.74	3.41 (1.61-7.23)	<0.01	5.22(1.95-13.97)	<0.01
2-4	BB, BA, AA (non-deficient)	245	304	392.27	0.77	1		1	
	AA- BA- (Heterozygous deficient)	69	81	111.72	0.73	0.94 (0.73-1.20)	0.59	0.96 (0.61-1.50)	0.86
	A-A- (deficient)	6	7	10.54	0.66	0.86 (0.41-1.81)	0.69	0.82(0.29-2.34)	0.72
4-6	BB, BA, AA (non-deficient)	228	205	343.59	0.60	1		1	
	AA- BA-	60	41	88.30	0.46	0.78 (0.56-1.09)	0.14	0.85 (0.53-1.36)	0.5

Age range	Genotypes	No. of participants (a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
	(Heterozygous deficient) A-A- (deficient)	7	6	11.98	0.50	0.84 (0.37-1.89)	0.67	0.54(0.25-1.16)	0.12
6-8	BB, BA, AA (non-deficient)	158	128	249.14	0.51	1		1	
	AA- BA-	45	33	69.73	0.47	0.92 (0.63-1.35)	0.67	0.89 (0.49-1.62)	0.71
	(Heterozygous deficient) A-A- (deficient)	7	4	11.63	0.34	0.67 (0.25-1.81)	0.43	0.52 (0.23-1.18)	0.12
8-10	BB, BA, AA (non-deficient)	94	46	133.44	0.34	1		1	
	AA- BA-	29	24	45.78	0.52	1.52 (0.93-2.49)	0.10	1.43 (0.77-2.65)	0.26
	(Heterozygous deficient) A-A- (deficient)	5	3	5.40	0.56	1.61 (0.50-5.18)	0.42	2.76(1.19-6.64)	0.02
>=10	BB, BA, AA (non-deficient)	42	8	47.35	0.17	1		1	
	AA- BA-	17	3	18.25	0	0.97 (0.26-3.67)	0.97	0.44(0.08-2.45)	0.35
	(Heterozygous deficient) A-A- (deficient)	2	0	2.46	0.00	_	_	_	_

4.4.5 CR1 McCoy

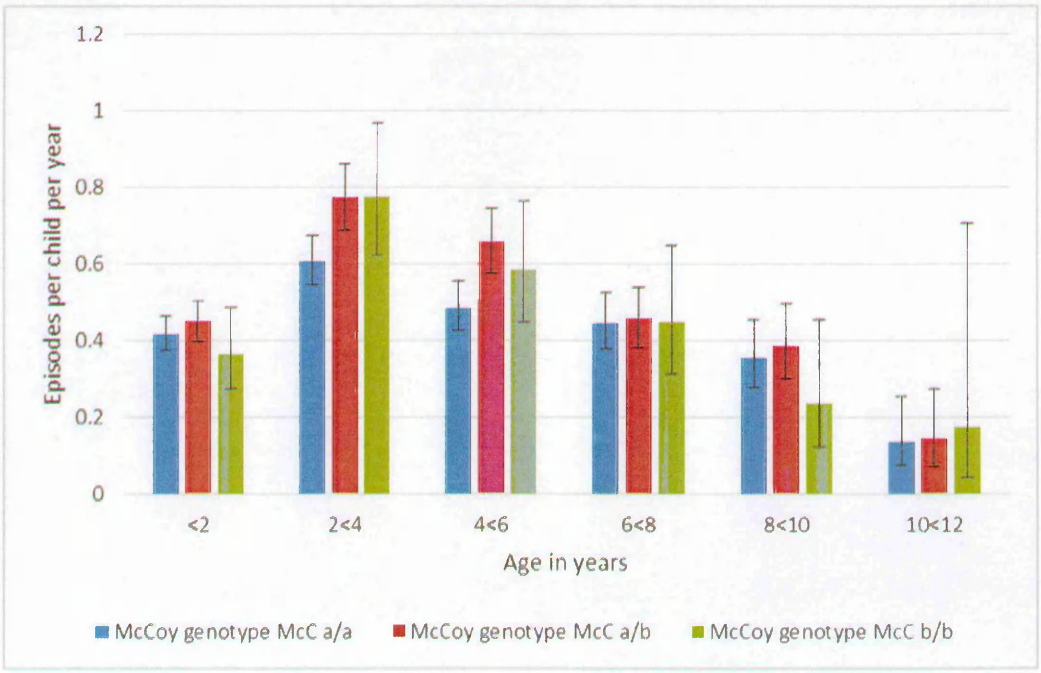
4.4.5.1 Effect of McCoy blood groups on incidence of malaria

McCoy genotypes consist of McC a/a, McC a/b and McC b/b. They result in McCoy blood groups. The effect of McCoy genotypes on malaria incidence is shown in Table 4.6. There was no significant effect of McCoy genotypes on malaria incidence. McCoy a/a was used as the reference group. McC a/b was associated with an IRR of 1.17 (0.98-1.40; p=0.08) whereas McC b/b individuals had IRR 1.04 (CI 0.82-1.32; p=0.74).

4.4.5.2 Age specific incidence of malaria by McCoy blood groups genotypes

McCoy rates were not sensitive to differences in age as shown in figure 4.6 and table 4.6 below.

Figure 4.6. Rates of uncomplicated malaria in McCoy genotypes stratified by age categories.



The y-axis represents the episodes of malaria per cyfu, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children with McC a/a genotype, red bars represent McC a/b genotype whereas green bars represent children McC b/b genotype.

The LRT comparing the interaction between McCoy genotypes and age with the non-genotype-age interaction model had a p-value of 0.17. There was no evidence that age interacts with McCoy genotypes to predict malaria incidence rates.

Table 4.6. Age-specific malaria incidence rates stratified by CR1 McCoy genotypes.

Age range	Genotypes	No. of participants(a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
All	Mc C a/a	705	1124	2409.15	0.47	1		1	
	Mc C a/b	577	1064	1958.40	0.54	1.16 (1.07-1.27)	<0.001	1.17 (0.98-1.40)	0.08
	Mc C b/b	113	222	439.16	0.51	1.08 (0.94-1.25)	0.27	1.04 (0.82-1.32)	0.74
0-2	Mc C a/a	586	345	824.54	0.42	1		1	
	Mc C a/b	460	295	655.69	0.45	1.08 (0.92-1.26)	0.36	1.10 (0.82-1.48)	0.51
	Mc C b/b	88	48	130.53	0.37	0.88 (0.65-1.19)	0.40	0.73 (0.50-1.07)	0.11
2-4	Mc C a/a	336	343	563.62	0.61	1		1	
	Mc C a/b	256	321	415.28	0.77	1.27 (1.09-1.48)	<0.01	1.30 (1.02-1.66)	0.03
	Mc C b/b	62	79	101.58	0.78	1.28 (1.00-1.63)	<0.05	1.24 (0.89-1.73)	0.21
4-6	Mc C a/a	312	218	447.54	0.49	1		1	
	Mc C a/b	238	241	366.84	0.66	1.35 (1.12-1.62)	<0.01	1.36 (1.00-1.85)	0.05
	Mc C b/b	57	55	93.71	0.59	1.20 (0.90-1.62)	0.22	1.13 (0.74-1.74)	0.56
6-8	Mc C a/a	208	144	321.61	0.45	1		1	
	Mc C a/b	178	135	295.53	0.46	1.02 0.81-1.29)	0.87	1.05 (0.74-1.49)	0.8
	Mc C b/b	42	29	64.18	0.45	1.01 (0.68-1.50)	0.96	1.29 (0.79-2.10)	0.3
8-10	Mc C a/a	118	64	179.55	0.36	1		1	
	Mc C a/b	117	63	162.33	0.39	1.09 (0.77-1.54)	0.63	0.96 (0.52-1.77)	0.89
	Mc C b/b	25	9	37.84	0.24	0.67 (0.33-1.34)	0.26	0.48 (0.15-1.49)	0.2
>=10	Mc C a/a	65	10	72.29	0.14	1		1	
	Mc C a/b	52	9	62.73	0.14	1.04 (0.42-2.55)	0.94	1.36 (0.46-4.00)	0.58
	Mc C b/b	11	2	11.32	0.18	1.28(0.28-5.83)	0.75	0.00	0.00

4.4.6 CR1 Swain Langley

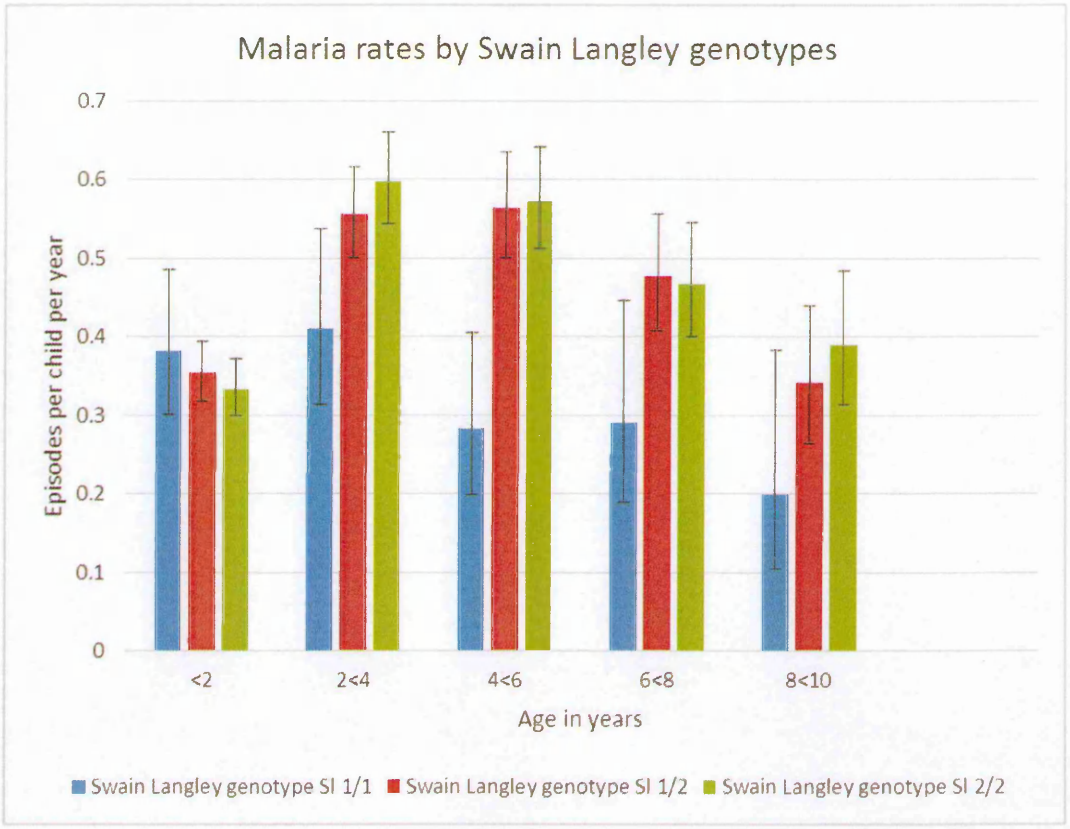
4.4.6.1 Effect of Swain-Langley blood groups on incidence of malaria

Sl genotypes can either be Sl 1/1 (normal), Sl 1/2 (heterozygotes) or Sl 2/2 homozygotes. Crude analyses indicated that Sl 1/2 and 2/2 variants increased the incidence of malaria. However, upon adjustment, the rates of malaria detected were the same with the reference group, Sl 1/1. The IRR for Sl 1/2 was 1.19 (0.86-1.64; p=0.29). Sl 2/2 had IRR 1.23 (0.90-1.70; p=0.20).

4.4.6.2 Age specific incidence of malaria by Swain-Langley genotypes

When the results were stratified by age, there was no difference with the pooled age results. The age stratified results are shown in figure 4.7 and table 4.7 below.

Figure 4.7. Rates of uncomplicated malaria in Swain Langley genotypes stratified by age categories.



The y-axis represents the episodes of malaria per cyfu, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children with SI 1/1 genotype, red bars represent SI 2/2 genotype whereas green bars represent children with SI 2/2 genotype.

The P value comparing the Swain Langley-age interaction model with the non-age interaction model was 0.19, indicating a non-significant interaction with age in Swain Langley genotypes. Age specific rates are tabulated in Table 4.7

Table 4.7. Age specific malaria incidence rates stratified by CR1 Swain Langley genotypes.

Age range	Genotypes	No. of participants(a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
(Years)									
All	Sl 1/1	128	162	431.72	0.38	1		1	
	Sl 1/2	624	1079	2161.34	0.50	1.33 (1.13-1.57)	<0.01	1.19 (0.86-1.64)	0.29
	Sl 2/2	645	1169	2216.00	0.53	1.40 (1.19-1.66)	<0.01	1.23 (0.90-1.70)	0.2
0-2	Sl 1/1	103	59	131.64	0.45	1		1	
	Sl 1/2	512	307	736.69	0.42	0.93 (0.70-1.23)	0.26	0.90 (0.55-1.46)	0.67
	Sl 2/2	521	322	744.78	0.43	0.96 (0.73-1.27)	0.1	0.88 (0.53-1.46)	0.63
2-4	Sl 1/1	54	43	89.68	0.48	1		1	
	Sl 1/2	303	332	503.08	0.66	1.38 (1.00-1.89)	0.13	1.29 (0.81-2.06)	0.28
	Sl 2/2	297	368	487.71	0.75	1.57 (1.15-2.16)	0.03	1.49 (0.93-2.38)	0.1
4-6	Sl 1/1	60	27	82.63	0.33	1		1	
	Sl 1/2	277	240	415.97	0.58	1.77 (1.19-2.63)	<0.01	1.68 (1.00-2.83)	0.05
	Sl 2/2	270	247	409.49	0.60	1.85 (1.24-2.75)	<0.01	1.61 (0.94-2.75)	0.08
6-8	Sl 1/1	42	21	66.94	0.31	1		1	
	Sl 1/2	192	141	303.25	0.46	1.48 (0.94-2.34)	0.06	1.30 (0.59-2.83)	0.51
	Sl 2/2	194	146	311.13	0.47	1.50 (0.95-2.36)	0.19	1.27 (0.59-2.74)	0.54
8-10	Sl 1/1	24	9	41.49	0.22	1		1	
	Sl 1/2	112	54	153.84	0.35	1.62 (0.80-3.28)	0.23	1.08 (0.36-3.26)	0.89
	Sl 2/2	124	73	184.39	0.40	1.82 (0.91-3.65)	0.17	1.12 (0.39-3.18)	0.83
>=10	Sl 1/1	17	3	19.33	0.16	1		1	
	Sl 1/2	47	5	48.51	0.10	0.66 (0.16-2.78)	0.97	1.04 (0.09-12.62)	0.98
	Sl 2/2	64	13	78.49	0.17	1.07 (0.30-3.75)	0.36	1.97 (0.24-16.10)	0.53

4.5 Does red blood cell genotype affect the parasitological case definition of malaria?

Using cross-sectional survey data (see section 2.5), the effect of genotypes on how parasite densities predict fever was investigated. This was done by conducting a likelihood ratio test comparing the logistic model with an interaction term between parasite density and genotype to a model without the genotype interaction term. If the two models are significantly different, then genotype affects the likelihood of having a malaria fever, and the parasite densities predicting fever should be reported in each genotype. The test was done for sickle, α-thalassaemia, ABO and G6PD genotypes and the results are reported in Table 3.11.

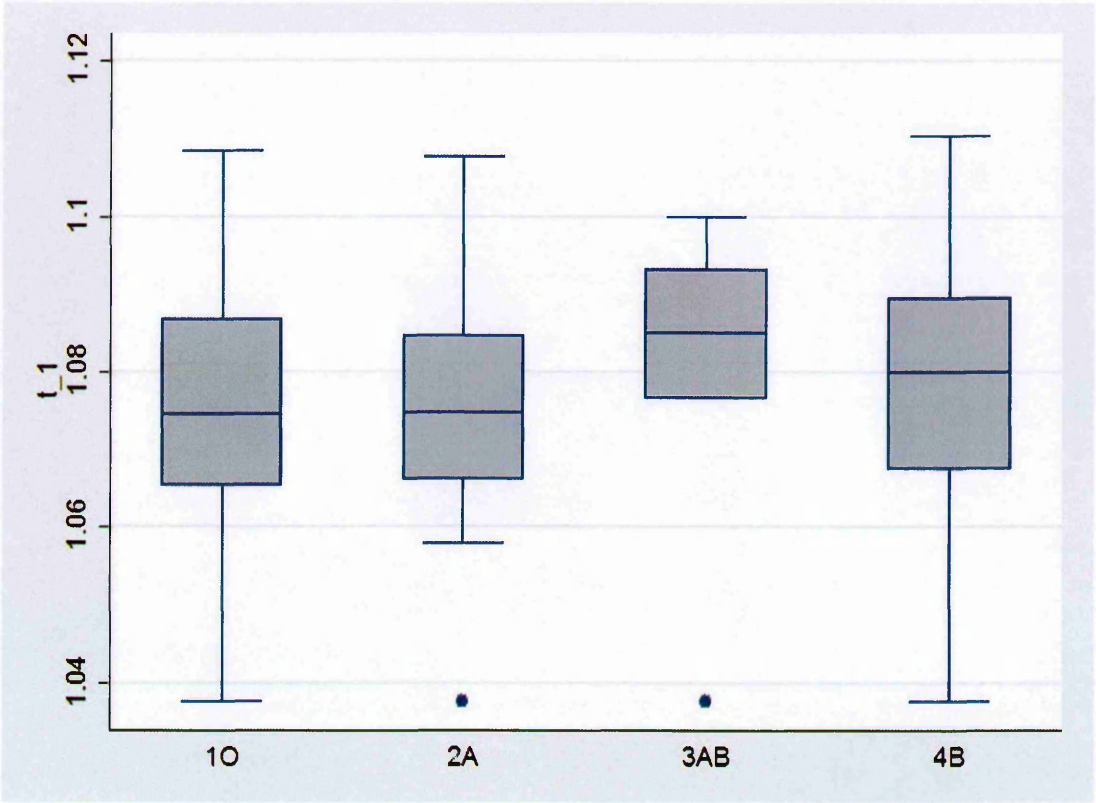
Table 4.8. Likelihood ratio test of red blood cell genotype interactions with parasite density in predicting fever compared with a model without interaction.

Interaction model	**LRT p-value**
Sickle genotype:	
model (i) logit fever ~ sickle_genotypes parasite density	
model (ii)logit fever ~ sickle_genotypes parasite density	
sickle_genotypes#parasite density	0.95
α+thalassaemia	
model (i) logit fever ~ α+thal_genotypes parasite density	
model (ii)logit fever ~ α+thal_genotypes parasite density	
α+thal_genotypes #parasite density	0.63
ABO blood group	
model (i) logit fever ~ abo_genotypes parasite_density	
model (ii) logit fever ~ abo_genotypes parasite_density	
abo_genotypes#parasite density	0.02
G6PD genotypes	
model (i) logit fever ~ g6pd_genotypes parasite_density	
model (ii)logit fever ~ g6pd_genotypes parasite_density	
g6pd_genotypes#parasite density	0.79

*G6PD genotype models were for the 202 genotype in males

There was no difference in the ability of the logistic model to predict a malaria fever based on interaction with parasite density for sickle, thalassaemia and G6PD genotypes. However, the ABO interaction model predicted fevers better than the non-interaction model. I hypothesized that this interaction effect may result from higher parasite densities in some blood groups of ABO. However, I found that the difference was not due to a difference in parasite densities in the ABO blood groups as shown in Figure 4.8.

Figure 4.8. Graph comparing parasite densities in ABO blood groups.



Key 1O=Blood group O 2A=Blood group A 3AB=Blood group AB 4B=Blood group B

t_1= best transformation of parasite density that predicts malaria fevers. For our model this transformation was (parasites per μL*0.01). There was no difference in parasite densities predicting a fever episode in the different blood groups.

4.6 Discussion and conclusions

Protective effects were shown in crude analysis of sickle cell trait, increased incidence of malaria in alpha thalassaemia homozygotes, increased incidence in non O blood groups, increased incidence in McCoy a/b blood groups and increased incidence in Sl 1/2 and Sl2/2 blood groups. However, adjustment for covariates and repeated episodes in the same individual led to a loss of effect. Using the LRT test, there was evidence suggesting that age interacts with some red blood cell genotypes such as sickle cell trait, α-thalassaemia, ABO blood groups and G6PD deficiency in males. However, there was no evidence of age and genotype interactions in predicting malaria rates among G6PD deficient females, McCoy and Swain Langley genotypes in males.

Age-specific incidence by genotypes and relation with naturally acquired immunity

Age as a proxy for malaria immunity: In many epidemiological studies of malaria, incidence varies with age, which is measure of exposure and therefore immunity to malaria. Age specific malaria incidence would therefore be a measure of immunity to malaria. This is a pragmatic approach especially because there is no reliable gold standard used as a measure of immunity. Including genotypes in this analysis showed whether genes affect rates of acquiring naturally acquired immunity to malaria, as described in a previous study (Williams, Mwangi, Roberts, *et al.*, 2005).

Age interactions affect naturally acquired immunity (Aponte *et al.*, 2007). It can be shown from malaria incidence rates stratified by genes which genes are protective from malaria infection. However, by conducting an age stratification analysis to show it was possible to

investigate if the effect is constant throughout the ages, or whether age affects the level of protection afforded by these genes.

Previous studies assessing evidence of acquired immunity include the study by Gong and colleagues (Gong *et al.*, 2012), who recruited a cohort of 601 children, 1-10 years of age in Kampala, Uganda, and followed them for 18 months for symptomatic malaria and asymptomatic parasitaemia. The molecular force of infection, a measure of establishment of parasitaemia was reduced in HbAS children 9 years of age (IRR 0.50; 0.28-0.87; P=0.01) compared to children at 2 years of age (IRR 1.16; 0.62-2.19; P=0.6). Similarly, protection from development of disease was more evident in older HbAS children at 9 years of age (RR 0.68; 0.51-0.91; P=0.008) compared to younger children at 2 years of age (RR 0.92; 0.77-1.10; P=0.3).

Le Hesran and colleagues (Le Hesran *et al.*, 1999) conducted a longitudinal study of *P. falciparum* infection and immunity. The prevalence of cellular responders to *P. falciparum* Pfl55/RESA antigen was similar in AA and AS children less than 2 years but was higher in AS children who were older. This implies that immunity to malaria develops at a higher rate among older children compared to infants.

If an immune basis to the protection exists, what might be its mechanism of action?

For HbAS

- Higher lympho-proliferative responses to *P. falciparum* soluble antigens in children with HbAS compared with normal children (Abu-Zeid *et al.*, 1992; Bayoumi, 1987).

- Higher titres of antibodies in HbAS compared with normal individuals(Marsh *et al.*, 1989).
- Interactive effect of HbAS in antibody-mediated immune protection (Achidi *et al.*, 1996).

For G6PD

- Courtin and colleagues (Courtin *et al.*, 2011) found that G6PD A- reduced the antibody response to *P. falciparum* antigens. However, it is still is still not clear if G6PD A- actually reduces the incidence of malaria in a prospective cohort.

Three important conclusions have resulted from this study. First, in the face of changing malaria epidemiology in the last few years, variants of red cell genes still provide varying levels protection from malaria incidence. Secondly, red cell genetic variants affect the rate of acquisition of naturally acquired immunity. The mechanisms through which red cell genes interact with immune genes to accelerate acquisition of naturally acquired immunity is an area requiring further investigations. Finally, the likelihood of having malaria fevers may be dependent on red cell genotypes, for example in ABO blood groups.

CHAPTER 5: The effect of other human candidate polymorphisms on malaria incidence

5.1 Introduction

Background: Apart from the classical red cell genetic polymorphisms, many other candidate malaria-protective polymorphisms have been proposed. In this chapter, I investigate the association between a range of additional SNPs in multiple genes and the incidence of uncomplicated malaria.

Methods: The methods used have been described in detail in the materials and methods section. In summary, DNA was extracted from blood samples contributed by 1462 study participants and typed for 67 SNPs in 38 genes using a multiplex Mass Array^{TM} method, as described in section 2.6.1.3 and Appendix IV. In addition samples were typed for α^{+}thalassaemia by PCR (as described in section 2.6.1.2) to give a total of 68 genetic variant sets in 39 genes.

Statistical analysis: The effect of the each polymorphism on malaria incidence was investigated by Poisson regression. The outcome variable was uncomplicated malaria incidence. Predictor variables included genotypes of the candidate genes and other covariates including age, gender, year, season, socio-economic status, the education level of the caregiver and distance to Lwak hospital. Each SNP was fitted with the covariates alone and the results reported. The results of a multivariate analysis including significant confounders are also reported.

Inferences: Since these SNPs are in malaria candidate genes, this investigation of their associations with incidence of malaria may be indicative of the role of the corresponding genes in malaria pathology. In addition, the results of these analyses are discussed in the context of the results from other studies that have investigated the effect on malaria of the same SNPs.

5.2 Objectives

1. To investigate allele frequencies and Hardy Weinberg equilibrium of SNPs in malaria candidate genes.
2. To investigate the effect of SNP markers in malaria candidate genes on incidence of malaria.

5.3 Results

5.3.1 Allele frequency, genotype frequency and Hardy Weinberg Equilibrium

Eleven SNPs were removed from further malaria incidence association analysis for the following reasons:

- Six were monomorphic in the study population: rs33930165, rs33950507, rs5743611, rs5743810, hcd36_g1439c and rs1799969.
- Two (rs2814778, rs1800482) had minor allele frequencies of less than 1%.
- Two alleles (rs461645 and rs7935564) deviated significantly ($p<0.01$) from Hardy Weinberg equilibrium.
- One allele, rs1800750 was not successfully genotyped by the multiplex assay.

As a quality control measure, 58 individuals were excluded because their self-reported gender was different from genetic markers in the amelogenin genes for gender (Eng *et al.*, 1994). Therefore, 56 SNPs were investigated for associations with uncomplicated malaria in 1462 individuals and the results are reported below. Deviations from Hardy Weinberg equilibrium were calculated using Pearson's χ^2 test. The allele frequencies and Hardy-Weinberg equilibrium tests are presented in Table 5.1 below.

Table 5. 1 Allele frequency and Hardy Weinberg Equilibrium for SNPs.

Gene name	rsnumber	Reference allele	Derived allele	p-value (Pearson's chi square)	Major allele	Minor allele	Minor allele frequency
Haemoglobin beta (HBB)	rs334	A	T	0.373	A	T	0.12
Guanylate binding protein 7	rs7537937	G	C	0.321	G	C	0.45
Interleukin 10	rs3024500	G	A	0.645	A	G	0.42
Interleukin 10	rs1800896	T	C	0.422	T	C	0.34
Interleukin 10	rs1800890	A	T	0.195	A	T	0.22
Complement Receptor 1 McCoy	rs17047660	A	G	0.738	A	G	0.28
Complement Receptor 1 Swain Langley	rs17047661	A	G	0.189	G	A	0.32
Interleukin 1A	rs17411697			0.389	G	T	0.17
Interleukin 1 B	rs1143634	G	A	0.254	C	T	0.10
Interleukin 17 RE	rs708567	C	T	0.683	A	G	0.46
Toll like receptor 9	rs352140	A	G	0.978	G	A	0.33
Toll like receptor 9	rs187084	G	A	0.996	T	C	0.30
Interleukin 17 RD	rs6780995	G	A	0.071	A	G	0.44
Toll like receptor 1	rs4833095	C	T	0.710	C	T	0.07
Toll like receptor 6	rs5743809	A	G	0.526	A	G	0.09
Complement 6	rs1801033	T	G	0.090	A	C	0.42
Interferon receptor 1	rs2706384	G	T	0.289	C	A	0.38
Interleukin 13	rs20541	G	A	0.694	C	T	0.22
Interleukin 4	rs2243250	C	T	0.439	T	C	0.22
Lymphotoxin alpha	rs2239704	C	A	0.054	G	T	0.17
Lymphotoxin alpha	rs909253	A	G	0.843	C	T	0.45

Gene name	rsnumber	Reference allele	Derived allele	p-value (Pearson's chi square)	Major allele	Minor allele	Minor allele frequency
Tumor Necrosis Factor	rs1799964	T	C	0.266	T	C	0.20
Tumor Necrois Factor	rs1800629			0.982	G	A	0.09
Tumor Necrosis Factor	rs361525	G	A	0.144	G	A	0.07
Tumor Necrosis Factor	rs3093662	A	G	0.497	A	G	0.12
Choline transporter-like protein 4	rs2242665	C	T	0.895	A	G	0.24
Interleukin 20 RA	rs1555498	C	T	0.949	C	T	0.45
Nucleotide-binding oligomerization domain-containing protein 1 (NOD1)	rs2075820	C	T	0.362	G	A	0.44
Cluster of differentiation 36 (CD36)	rs3211938	T	G	0.682	T	G	0.08
Cystic fibrosis transmembrane conductance regulator (CFTR)	rs17140229	T	C	0.136	T	C	0.29
Toll-like receptor 4 (TLR4)	rs4986790	A	G	0.984	A	G	0.08
Toll-like receptor 4 (TLR4)	rs4986791	C	T	0.819	C	T	0.01
Reticulon-3 (RTN3)	rs542998	T	C	0.386	T	C	0.39
Interleukin 22 (IL22)	rs2227507	T	C	0.144	T	C	0.03
Interleukin 22(IL22)	rs1012356	A	T	0.977	A	T	0.50
Interleukin 22(IL22)	rs2227491	G	A	0.390	C	T	0.40
Interleukin 22(IL22)	rs2227485	G	A	0.333	G	A	0.39
Interleukin 22(IL22)	rs2227478	G	A	0.818	A	G	0.41
Spectrin beta chain, erythrocyte (SPTB)	rs229587	T	C	0.588	T	C	0.34
Adenylate cyclase type 9 (ADCY9)	rs2230739	T	C	0.912	A	G	0.07

Gene name	rsnumber	Reference allele	Derived allele	p-value (Pearson's chi square)	Major allele	Minor allele	Minor allele frequency
Adenylate cyclase type 9 (ADCY9)	rs10775349	C	G	0.978	C	G	0.22
interleukin 4 receptor (IL4R)	rs1805015	T	C	0.422	T	C	0.46
Adenosine A2b receptor (ADORA2B)	rs2535611	C	T	0.678	T	C	0.11
Nitric oxide synthase 2A (NOS2A)	rs2297518	G	C	0.514	G	A	0.12
Nitric oxide synthase 2A (NOS2A)	rs9282799	G	A	0.048	C	T	0.05
Nitric oxide synthase 2A (NOS2A)	rs8078340	G	A	0.825	C	T	0.19
EGF-like module-containing mucin-like hormone receptor-like 1 (EMR1)	rs373533	C	A	0.266	G	T	0.47
Intercellular adhesion molecule 1 (ICAM1)	rs5498	A	G	0.951	A	G	0.09
Guanine Nucleotide Binding Protein (G Protein), Alpha Stimulating Activity Polypeptide 1 (GNAS)	rs8386	C	T	0.294	C	T	0.15
Derlin-3 (DERL3)	rs3177244	G	A	0.751	A	G	0.49
Haemoglobin alpha (HBA)	NA	Normal	3.7kb Deletion	0.375	Normal	Deletion	0.32

5.3.2 Association of SNPs with uncomplicated malaria

The association was investigated by determining a model and significance threshold as defined in section 5.3.2.1. The effect of SNPs on autosomal genes is reported in section 5.3.2.2 whereas the effect of SNPs on X linked genes is reported in section 5.3.2.3.

5.3.2.1 Model description and determination of a significance threshold

Associations between SNPs and malaria were studied using a Poisson model with the covariates age, gender, education, socio-economic status, year, season and distance to Lwak hospital (km). All the SNPs that yielded a p-value of <0.1 were fitted into a multi-gene SNP model with covariates as listed above. Since multiple statistical testing inflates the occurrence of false positives, an adjustment for multiple testing was required. Bonferroni correction was used to adjust for multiple testing by dividing the conventional type I error p-value (0.05) by the number of SNPs tested. However, since Bonferroni correction is a conservative adjustment method, a number of promising SNPs did not attain the statistical significance cut-off. Some of these will be discussed in the results as gene markers that potentially influence malaria incidence.

5.3.2.2 Autosomal genes: Effect of 50 SNPs on malaria incidence

Each gene was tested individually for its malaria protective effect, with and without adjustment for the confounding variables such as the distance to hospital, age, year, season, education, socio-economic status and gender. Significant results of at least one gene variant are presented in Table 5.2 below.

Table 5. 2 Genotypes that significantly affect the incidence of malaria in Poisson regression analyses.

rsnumber	Gene_name	genotype	IRR	min95	max95	p_value
rs17411697	Interleukin 1A	TT	2.19	1.31	3.66	<0.01
rs352140	Toll like receptor 9	AG	1.19	1	1.42	0.05
rs5743809	Toll like receptor 6	CC	0.26	0.15	0.45	<0.0001
rs1801033	Complement 6	AC	0.68	0.53	0.88	<0.01
rs1801033	Complement 6	CC	0.75	0.58	0.98	0.04
rs20541	Interleukin 13	TT	1.67	1.16	2.41	0.01
rs2243250	Interleukin 4	CT	1.45	1.01	2.08	0.04
rs2243250	Interleukin 4	CC	1.56	1.09	2.21	0.01
rs1800629	Tumor Necrosis Factor	AG	1.24	1	1.53	0.05
rs361525	Tumor Necrosis Factor	AA	0.49	0.29	0.84	0.01
rs3093662	Tumor Necrosis Factor	GG	0.56	0.36	0.86	0.01
rs17140229	CFTR	CT	1.25	1.05	1.49	0.01
rs4986790	TLR4	GG	0.28	0.16	0.51	<0.0001
rs7935564	TRIM5	AG	0.83	0.69	0.99	0.04
rs2227507	IL22	CC	0	0	0	<0.0001
rs2230739	ADCY9	GG	3.59	1.88	6.82	<0.0001
rs1805015	IL4R	CT	1.25	1.03	1.52	0.02
rs8078340	NOS2A	TT	1.58	1.01	2.46	0.04
rs373533	EMR1	TT	0.77	0.61	0.97	0.02

Considering only the above genes, a simple model of innate immunity may be that pathogens are recognized through toll like receptors, innate immune effector mechanisms are recruited into action through production of TNF and interleukins, this process may be affected by cell signalling and membrane receptors and finally infected cells are killed by the complement system or through the action of species such as NO and TRIM-5. The presence of mutations affecting the incidence of malaria along this pathway are discussed below.

5.3.2.2.1 Toll like receptors affect uncomplicated malaria incidence

(i) TLR-9 rs352140 G>A mutation is associated with an increased incidence of uncomplicated malaria.

Toll like receptors are innate immune receptors that recognize pathogen associated molecular patterns and trigger activation of immune cell responses. The current study used GG genotype of the rs352140 SNP as the reference group to compare malaria incidence across genotypes. Heterozygosity for AG was associated with a higher incidence of uncomplicated malaria (IRR 1.19; 1.00-1.42; $p<0.05$ but no significant association was seen with homozygosity of the A allele (IRR 1.15; 0.84-1.57; $p=0.39$). In keeping with these findings, TLR9 rs352140 GG was associated with a reduced risk of symptomatic malaria (RR 0.34; $p=0.048$) in a cohort study conducted in Ghanaian children (Omar *et al.*, 2012). This mutation has also been associated with gene interactions which control parasite infection levels (Basu *et al.*, 2012).

(ii) Toll like receptor 6 rs5743809 T>C mutation reduces the incidence of uncomplicated malaria

Using the TT genotype as the reference, homozygosity for the C allele was associated with a lower incidence of uncomplicated malaria (IRR 0.26; 0.15-0.45; $p<0.0001$). No association was seen with heterozygosity (IRR 0.94; 0.0.76-1.16; $p=0.54$). 68% of these genotypes were successfully genotyped. However, this finding is similar to that by Leorratti and colleagues, who reported that the TT homozygotes were over represented in individuals with asymptomatic infections (Leoratti *et al.*, 2008). Although the function of the rs5743809 T>C polymorphism is unknown, TLR 6 forms heterodimers with TLR2, a molecule that recognizes *P. falciparum* glycosylphosphatidylinositol (GPI) to recruit the innate immune

system to eliminate the parasite. It is likely that a disruption of any of these molecules affects the role of TLR2 in clearance of the parasite from the system (Leoratti *et al.*, 2008).

iii) TLR-4 rs4986790 A>G mutation reduces the incidence of uncomplicated malaria

TLR-4 also recognises *P. falciparum* glycosylphosphatidylinositol, and primes the innate immunity for clearance of the parasites from the body (Krishnegowda *et al.*, 2005). The TLR-4 gene carries the rs4986790 SNP also known as Asp299Gly (or 896A/G). Compared to the AA genotype, incidence of malaria was significantly reduced in GG homozygotes (IRR 0.28; 0.16-0.51; p<0.0001) but not in AG heterozygotes (IRR 0.89; 0.71-1.1; p=0.27). The GG homozygote frequency was 1%, therefore the power was low to detect an effect. In a study in Burundi, genotype frequencies were similar between severe and uncomplicated malaria cases (Esposito *et al.*, 2012).

5.3.2.2.2 Interleukin gene mutations and their receptors affect uncomplicated malaria incidence

i) Interleukin-1 α rs17411697 G>T mutation increases the incidence of uncomplicated malaria.

This SNP in IL-1 α is also known in other notations as rs17561. It codes for a non-conservative amino acid change from alanine to serine at codon 114 of the interleukin 1-α protein. In this study, the GG genotype was used as the reference group. Homozygosity for the T allele was associated with raised incidence of malaria (IRR 2.19; 1.31-3.66; p <0.01). No effect was seen in GT heterozygotes (IRR 1.13; 0.95-1.34; p=0.17). My findings are similar to a previous study reporting an association between this polymorphism and uncomplicated malaria in Gambian children (Walley *et al.*, 2004). However, a second study

has reported an association between this SNP and severe malaria in Vietnamese adults (Dunstan *et al.*, 2012). As discussed by Dunstan and colleagues (Dunstan *et al.*, 2012) the disparity in results might be explained by differences in case definitions, demographics of the malarial populations or to differences in transmission. At the time the Gambian study was conducted, The Gambia was a high transmission area, where most of the population had acquired immunity to malaria and would be expected to show only mild symptoms of the disease. As in most immune populations, clinical disease commonly affects children who have not fully developed immunity to malaria. However, in Vietnam low transmission results in a non-immune population, resulting in the disease affecting all age categories, and therefore disease presentation may be severe in the affected adults. The finding of association with mild malarial disease does not preclude association with severe disease. In the current study, this association could not be tested due the low number of severe disease episodes.

ii) IL-13: rs20541 C>T mutation increases the incidence of uncomplicated malaria.

Interleukin 13 is a Th2 cytokine that has about 30% sequence similarity with IL-4 and has similar functions (Zurawski & de Vries, 1994). It activates B cells and monocytes, promotes IgE class switching and inhibits the production of pro-inflammatory cytokines. In the current study, CC genotype individuals were the reference group. Homozygosity for the T allele (IRR 1.67; 1.16-2.41; $p< 0.01$) was associated with higher incidence of malaria. Heterozygous CT individuals (IRR 0.95; 0.80-1.13; $p=0.54$) did not show a difference compared to the reference group. The C allele frequency was 0.78 whereas the T allele frequency was 0.22. The role of the mutation is well defined for severe disease compared to uncomplicated malaria. In a case control study conducted among Vietnamese, this mutation was associated with protection from severe malaria (Dunstan *et al.*, 2012). A mutation in the

same locus, the IL13 promoter -1055 C to T has also been associated with severe malaria protection in a case control study in Thailand (Ohashi *et al.*, 2003). More recently, a Tanzanian case-control study found that the rs20541 TT children were twice as likely to develop cerebral malaria compared to the CT children (Manjurano *et al.*, 2012). Taken together, the higher incidence of uncomplicated malaria in the TT genotype reported in this study and more severe disease in the TT genotype from previous studies are consistent with a role of IL-13 in susceptibility to mild malaria and progression from mild to severe malarial disease.

iii) IL-4 rs2243250 T>C mutation increases the incidence of uncomplicated malaria.

Interleukin-4 is a cytokine that activates T-cells from naïve forms to T-helper 2 (Th2) cells. Th2 cells are important host immune effectors which stimulate proliferation of B-cells, increase production of neutralizing antibodies and induce the switch between antibody classes. The IL-4 rs2243250T variant is associated with increased transcriptional activity of the IL4 gene. Conversely, the C variant would be associated with lower levels of IL-4. Using the TT as the reference, both the CT heterozygotes (IRR 1.45; 1.01-2.08; p=0.04) and CC homozygotes (IRR 1.56; 1.09-2.21; p<0.01) had a significantly higher incidence of mild malaria. However, the relationship between this variant and malaria is not the same between different studies. In family based pedigree studies in Tanzania, no association was reported between this locus and clinical episodes of malaria (Carpenter *et al.*, 2009). A study conducted in the Burkina Faso found that the mutation was not associated with *P. falciparum*-specific immunoglobulin E (IgE) in severe malaria or uncomplicated malaria. However, total IgE in IL-4 rs2243250 T individuals was significantly elevated among children with severe malaria (Verra *et al.*, 2004). The authors proposed there could be an association between IgE

and severe malaria, and that this relationship may be investigated in other settings. Another study in Ghanaian children with cerebral malaria found that the T variant appears at higher frequency in children with cerebral malaria compared with children with mild malaria (Gyan *et al.*, 2004). Since cerebral malaria normally manifests in children with lower exposure to malaria, it would appear that the C variant protects from severe disease by predisposing children to higher mild malaria disease episodes.

iv) IL4R rs1805015 T>C mutation increases the incidence of uncomplicated malaria.

The importance of IL-4 in malaria incidence is further reinforced by the fact that the IL4 receptor was also associated with malaria incidence. Compared to the TT genotype, incidence of malaria was significantly reduced in CT heterozygotes (IRR 1.25; 1.03-1.52; $p<0.02$) but not in CC homozygotes (IRR 1.14; 0.88-1.47; $p=0.31$). A study in Gabonese patients with severe malaria found high levels of soluble IL-4 receptor in serum (Perkmann *et al.*, 2005), suggesting a role of interleukin 4 in severe disease.

v) IL-22 rs2227507 T>C mutation may reduce the incidence of uncomplicated malaria in homozygotes.

IL-22 modulates tissue responses during inflammation, and is produced in response to IL-12, IL-18 and IL-13 by natural killer cells (Zenewicz & Flavell, 2011). The current study used the TT genotype of rs2227507 as the reference, and found that the incidence of malaria appeared to be reduced in the rare CC homozygotes ($p<0.0001$). However, there were only 3 CC homozygotes (genotype frequency < 1%) and therefore the data could not allow for an estimate of the IRR or confidence intervals. As such, the result should be interpreted with caution. There was no effect of AG heterozygosity (IRR 1.15; 0.95-1.4; $p=0.16$). A study in West Africa identified two haplotypes in IL22, one containing the rs2227491 (+708) that was

associated with protection against severe anaemia and another containing the rs2227478 (-1394) that was weakly associated with cerebral malaria. Since these SNPs could exist in different haplotype permutations with the one in the current study, the haplotypes need to be further investigated to delineate the causal allele for the reduced incidence reported in this study.

5.3.2.2.3 Effect of polymorphisms in TNF on malaria incidence

i) TNF rs1800629 G>A mutation is protective against the incidence of uncomplicated malaria.

TNF is a cytotoxin, produced by monocytes as part of the innate immune response. Compared to the GG genotype, incidence of malaria was significantly reduced in AG heterozygotes (IRR 1.24; 1.00-1.53; $p<0.05$) but not in AA homozygotes (IRR 0.88; 0.43-1.8; $p=0.73$). Higher plasma levels of TNF have been linked with cerebral malaria and malarial fatality (Kwiatkowski *et al.*, 1990). In another study conducted in the Solomon Islands, serum TNF levels were found to be correlated with malarial parasite density and patient's temperatures (Butcher *et al.*, 1990). Thus SNPs which may affect TNF function or production are expected to have an effect on malaria symptoms.

ii) TNF rs361525 G>A mutation reduces the incidence of uncomplicated malaria.

Compared to the GG genotype, incidence of malaria was significantly reduced in AA homozygotes (IRR 0.49; 0.29-0.84; $p<0.01$) but not in AG heterozygotes (IRR 1.08; 0.84-1.4; $p=0.54$). However, the significant reduction in AA homozygotes should be interpreted with caution as the AA genotype frequency was only 1%.

iii) TNF rs3093662 A>G mutation reduces the incidence of uncomplicated malaria.

Compared to the AA genotype, incidence of malaria was significantly reduced in GG homozygotes (IRR 0.56; 0.36-0.86; p<0.01) but not in AG heterozygotes (IRR 1.15; 0.95-1.4; p=0.16). The GG homozygote frequency was 2%, therefore the result should be interpreted with caution.

5.3.2.2.4 Mutations in membrane receptors and cell signalling pathway receptors affect mild malaria incidence

i) EMR1 rs373533 G>T mutation is protective against incidence of uncomplicated malaria.

Epidermal growth factor module-containing mucin-like hormone receptor (EMR) 1 is a hormone receptor which belongs to the 7 trans-membrane superfamily of receptors. Compared to the GG genotype, incidence of malaria was significantly reduced in TT homozygotes (IRR 0.77; 0.61-0.97; p=0.02) but not in GT heterozygotes (IRR 0.92; 0.75-1.12; p=0.4). The rs373533 mutation has been associated with malaria-associated seizures in a case control study conducted in Kumasi (Kariuki *et al.*, 2013). In a Cameroonian study of patients admitted with malaria, the GT individuals were found to have a significantly higher risk of hyperpyrexia compared to the GG and TT individuals (Apinjoh *et al.*, 2014).

ii) CFTR rs17140229 T>C mutation increases the incidence of uncomplicated malaria.

CFTR is a chloride channel in epithelial cell membranes. Mutations in CFTR have previously been associated with protection from *Salmonella typhi* infection (van de Vosse *et al.*, 2005). The current study used the TT genotype as the reference group and found that the incidence of malaria was significantly increased in CT heterozygotes (IRR 1.25; 1.05-1.49; p<0.01) but not in CC homozygotes (IRR 1.31; 0.96-1.78; p=0.08). A previous study in Cameroon comparing the TT with the CT and CC variants has associated the latter 2 variants with a

protective effect from malarial anaemia in Cameroonian patients (Apinjoh *et al.*, 2014). This SNP has also been associated with repetitive malaria associated seizures in a pooled analysis of data from Kenya, Malawi, Tanzania and Ghana (Kariuki *et al.*, 2013).

iii) ADCY9 rs2230739 A>G mutation may increases the incidence of uncomplicated malaria.

Adenylate cyclase 9 is a stimulatory G protein signal transduction protein that is membrane bound. Compared to the AA genotype, incidence of malaria was significantly increased in GG homozygotes (IRR 3.59; 1.88-6.82, $p<0.0001$) but not in AG heterozygotes (IRR 0.89; 0.68-1.17; $p=0.41$). The GG homozygote frequency was 0.4%, therefore the result should be interpreted with caution. However, Auburn and colleagues did not identify any association between this allele and severe malaria (Auburn *et al.*, 2010).

5.3.2.2.5 Mutations in genes involved in clearing pathogens from the body affect mild malaria incidence

i) Complement 6 rs1801033 A>C mutation is protective from incidence of uncomplicated malaria.

Complement 6 is part of the innate complement system, involved in the classical pathway of complement activation, which complements the ability of phagocytic cells and antibodies to clear foreign antigens from the body. This protein constitutes a part of the membrane attack complex, which usually causes cells that are infected with pathogens to lyse. Using the AA genotype as the reference group, both AC heterozygotes (IRR 0.68; 0.53-0.88; $p<0.01$) and CC homozygotes (IRR 0.75; 0.58-0.98; $p=0.04$) were found to have a lower incidence of uncomplicated malaria. The rs1801033 A>C SNP may lead to more efficient action of

complement 6 and therefore lead to less occurrence of clinical malaria in individuals carrying this mutation, as observed in this study.

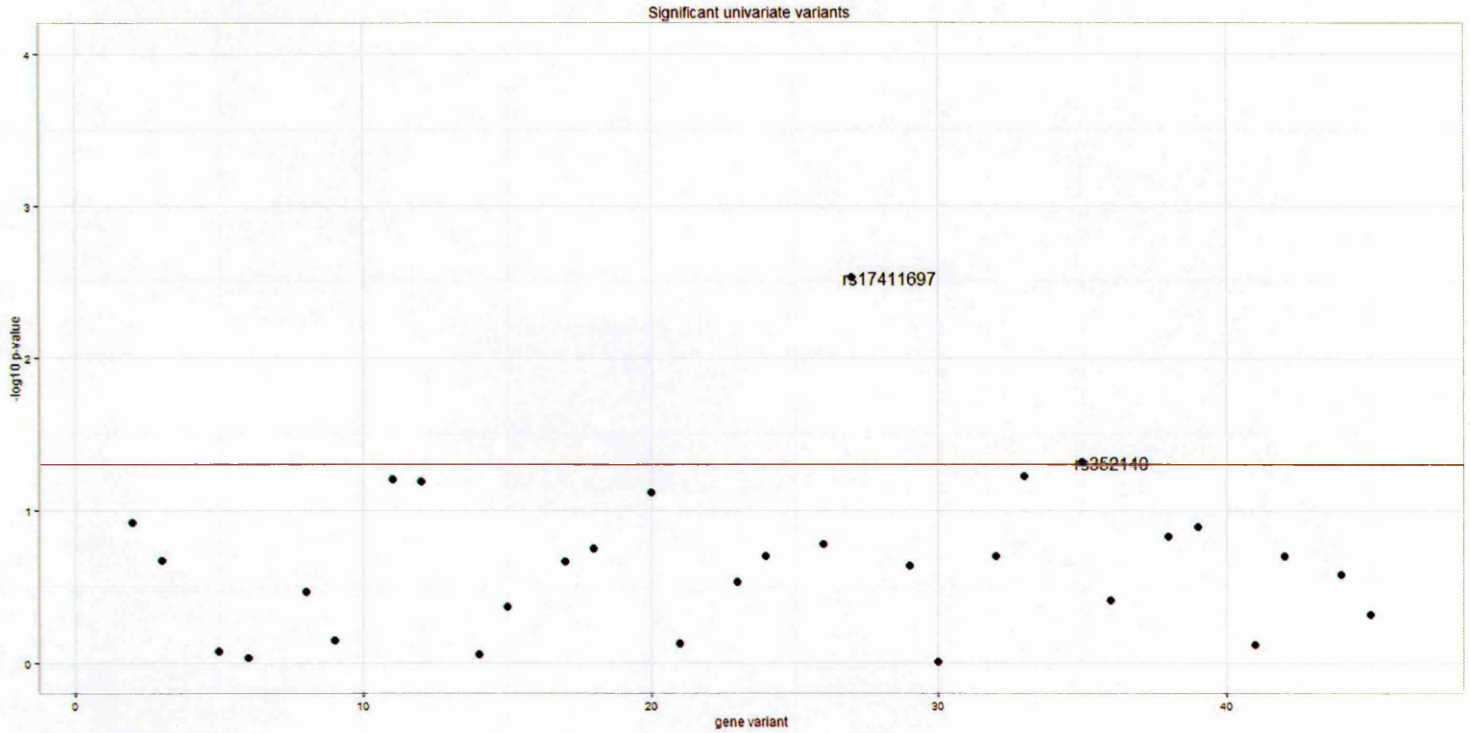
ii) NOS2A rs8078340 C>T mutation increases the incidence of uncomplicated malaria. Compared to the CC genotype, incidence of malaria was significantly reduced in TT homozygotes (IRR 1.58; 1.01-2.46; p=0.04) but not in CT heterozygotes (IRR 0.99; 0.83-1.19; p=0.96). The TT homozygote frequency was 4%, therefore the result should be interpreted with caution. In contrast, a study in Vietnam (Dunstan *et al.*, 2012) did not find a polymorphism at rs8078340. However, nitric oxide species have previously been reported to have an inverse relationship to severe malaria outcomes (Anstey *et al.*, 1996), with higher levels in asymptomatic individuals and the lowest levels in children who had cerebral malaria. Mutations in genes such as NOS2A involved in production of nitric oxide are therefore expected to have an effect on malaria outcomes.

iii) TRIM-5 rs7935564 G>A mutation reduces the incidence of uncomplicated malaria. Mutations in TRIM 5 have been associated with susceptibility to HIV (Nakayama & Shioda, 2012) possibly because the molecule binds to HIV virions and targets them for removal from the body (Lukic *et al.*, 2011). Compared to the GG genotype, incidence of malaria was significantly reduced in AG heterozygotes (IRR 0.83; 0.69-0.99; p=0.04) but not in AA homozygotes (IRR 1.16; 0.88-1.54; p=0.3. This SNP has previously been recognized as a risk factor for severe malaria (additive A, OR= 1.273; 1.042–1.555; p=0.02) in Tanzania (Manjurano *et al.*, 2012). In a study in Vietnam, this SNP was associated with death from severe malaria when the severe malaria episodes were compared with cases that recovered from the disease (Dunstan *et al.*, 2012).

Full results for all variants of the genes tested are presented in Figure 5.1 (a)-(d) and Table 5.3 below.

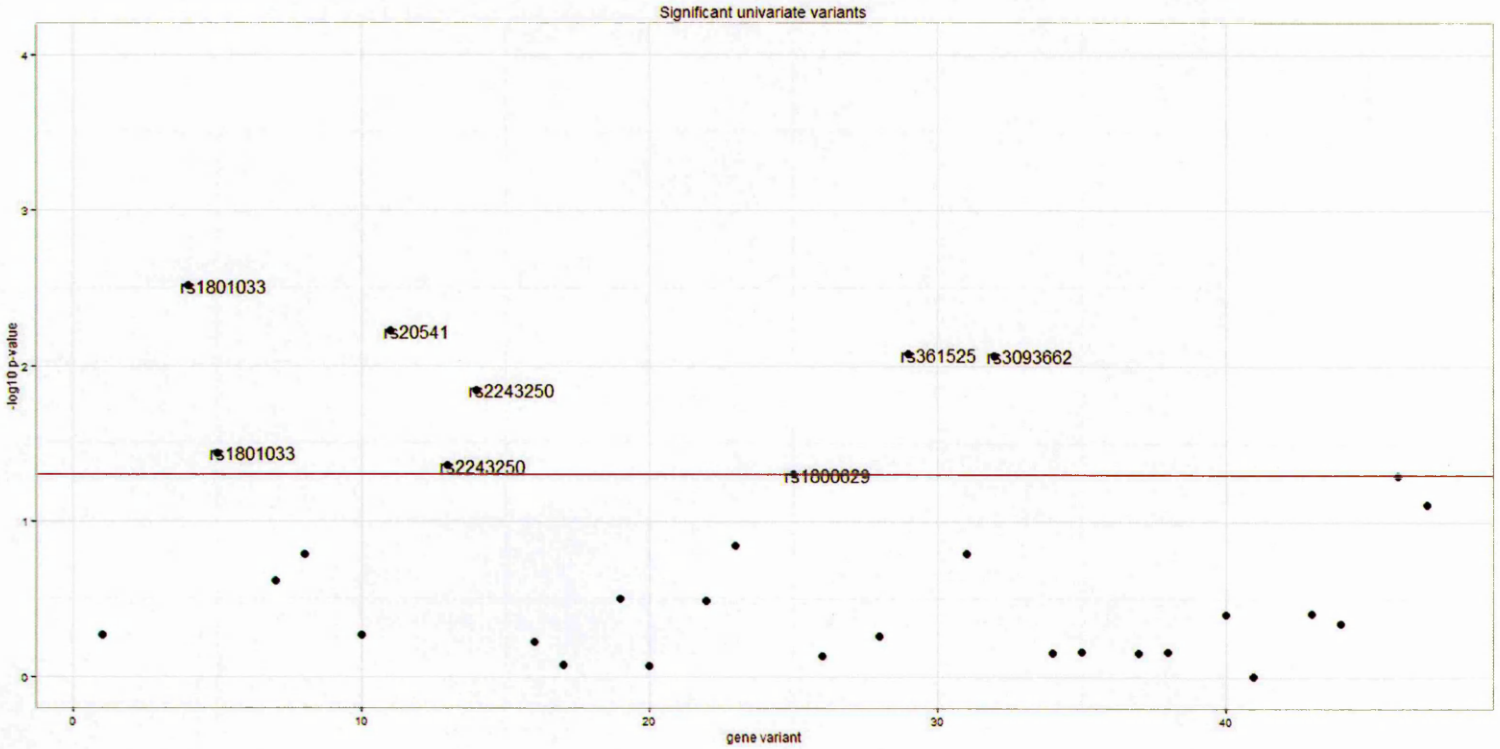
Figure 5. 1 Manhattan plots showing associations of the SNPs with uncomplicated malaria.

Part (a) Gene variants 1-47



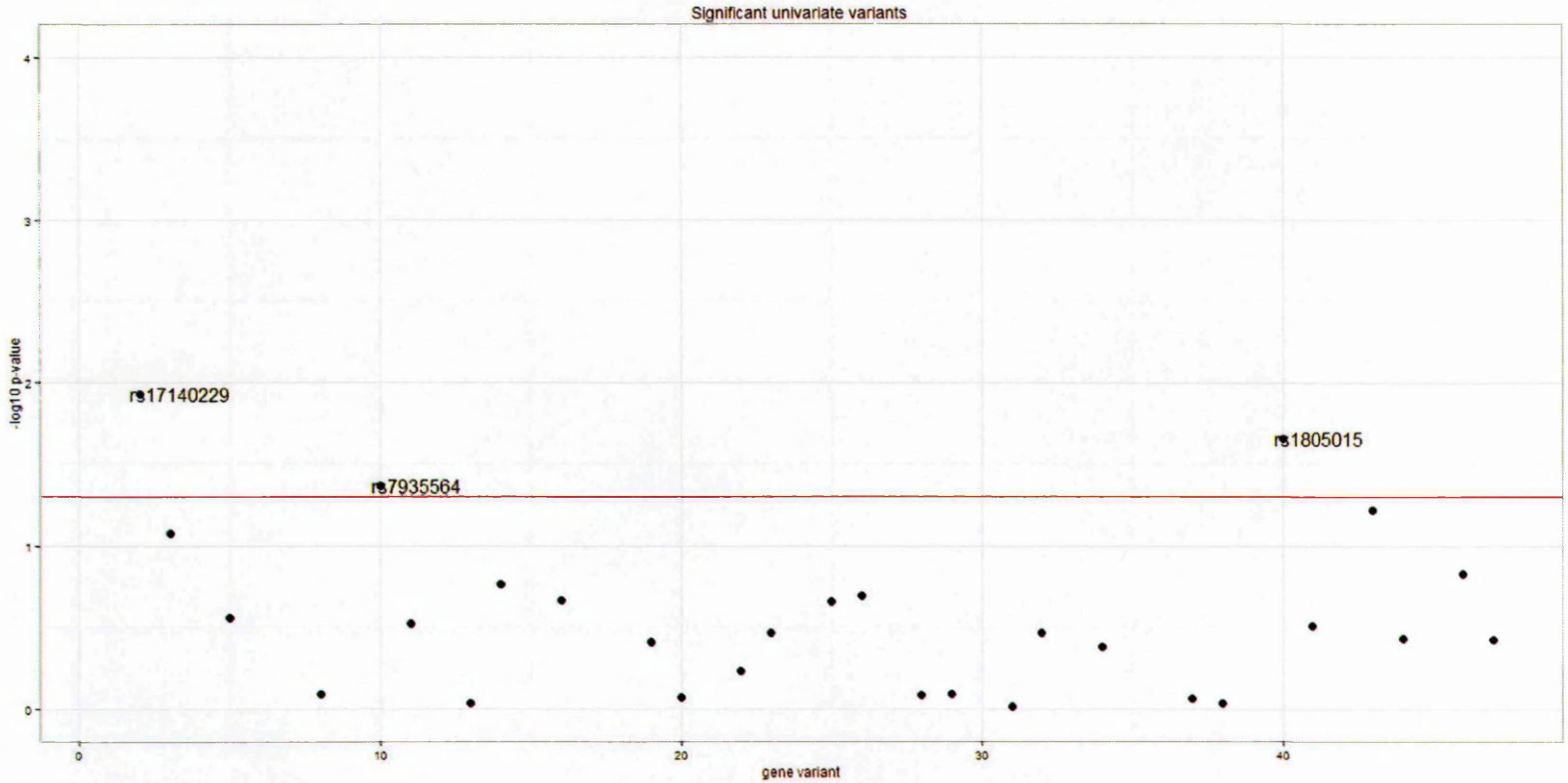
Each SNP is plotted for the homozygous, heterozygous and normal variants. SNP variants that significantly affect malaria incidence appear above the red line.

Part b) Gene variants 48-94



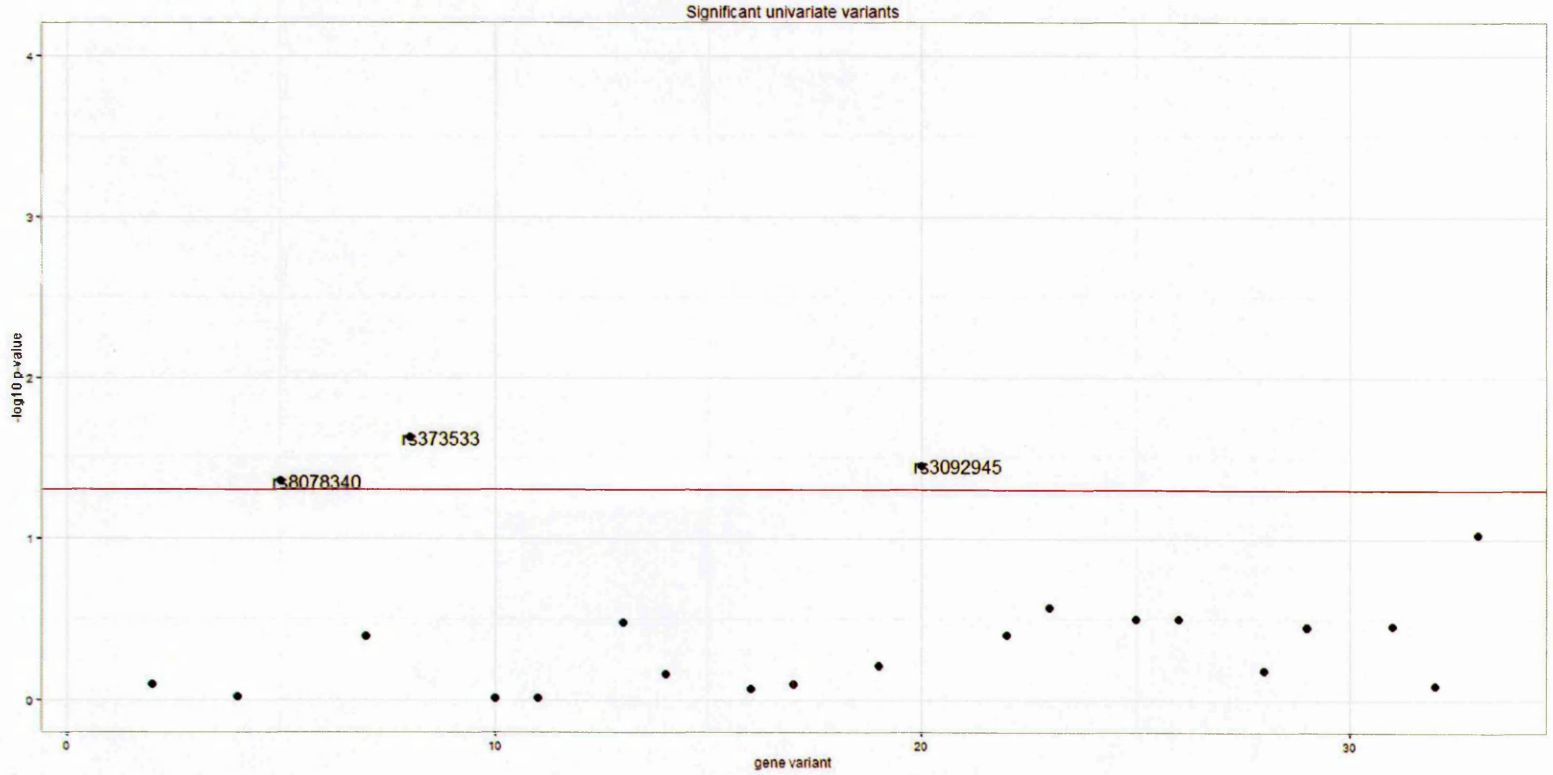
Each SNP is plotted for the homozygous, heterozygous and normal variants. SNP variants that significantly affect malaria incidence appear above the red line.

Part c) Gene variants 95-141



Each SNP is plotted for the homozygous, heterozygous and normal variants. SNP variants that significantly affect malaria incidence appear above the red line.

Part d) Gene variants 142-173



Each SNP is plotted for the homozygous, heterozygous and normal variants. SNP variants that significantly affect malaria incidence appear above the red line.

Table 5. 3 The effect of SNPs on uncomplicated malaria incidence rates, controlled for covariates

Gene name	rs number	Genotype	Number of participants	IRR	LCL	UCL	p_value
GBP7	rs7537937	GG	423	1.00	NA	NA	
		CG	673	0.91	0.75	1.1	0.34
		CC	298	1.05	0.83	1.32	0.71
Interleukin 10	rs3024500	AA	246	1.00	NA	NA	
		AG	690	1.25	0.99	1.57	0.06
		GG	460	1.28	0.98	1.67	0.06
Interleukin 10	rs1800896	TT	608	1.00	NA	NA	
		CT	634	0.99	0.83	1.17	0.88
		CC	150	0.88	0.64	1.21	0.43
Interleukin 10	rs1800890	AA	839	1.00	NA	NA	
		AT	498	0.90	0.76	1.06	0.21
		TT	60	0.64	0.34	1.22	0.18
Interleukin 1A	rs17411697	GG	957	1.00	NA	NA	
		GT	404	1.13	0.95	1.34	0.17
		TT	36	2.19	1.31	3.66	<0.01
Interleukin 1 B	rs1143634	CC	1133	1.00	NA	NA	
		CT	245	1.14	0.92	1.4	0.23
		TT	18	1.01	0.54	1.89	0.98
Interleukin 17 RE	rs708567	AA	288	1.00	NA	NA	
		AG	700	0.87	0.71	1.07	0.2
		GG	407	0.80	0.63	1.01	0.06
Toll like receptor 9	rs352140	GG	624	1.00	NA	NA	
		AG	619	1.19	1.00	1.42	<0.05
		AA	153	1.15	0.84	1.57	0.39
Toll like receptor 9	rs187084	TT	127	1.00	NA	NA	
		CT	588	1.27	0.92	1.75	0.15
		CC	681	1.28	0.93	1.76	0.13

Gene name	rs number	Genotype	Number of participants	IRR	LCL	UCL	p_value
Interleukin 17 RD	rs6780995	AA	248	1.00	NA	NA	
		AG	719	1.04	0.82	1.32	0.75
		GG	428	1.18	0.92	1.5	0.2
Toll like receptor 1	rs4833095	CC	1174	1.00	NA	NA	
		CT	182	1.15	0.90	1.46	0.26
		TT	6	0.67	0.22	2.03	0.48
Toll like receptor 6	rs5743809	TT	1152	1.00	NA	NA	
		CT	231	0.94	0.76	1.16	0.54
		CC	14	0.26	0.15	0.45	<0.0001
Complement 6	rs1801033	AA	227	1.00	NA	NA	
		AC	709	0.68	0.53	0.88	<0.01
		CC	459	0.75	0.58	0.98	0.04
Interferon receptor 1	rs2706384	CC	194	1.00	NA	NA	
		AC	678	0.86	0.67	1.11	0.24
		AA	525	0.83	0.64	1.08	0.16
Interleukin 13	rs20541	CC	844	1.00	NA	NA	
		CT	475	0.95	0.80	1.13	0.54
		TT	71	1.67	1.16	2.41	<0.01
Interleukin 4	rs2243250	TT	71	1.00	NA	NA	
		CT	457	1.45	1.01	2.08	0.04
		CC	829	1.56	1.09	2.21	<0.01
Lymphotoxin alpha	rs2239704	GG	975	1.00	NA	NA	
		GT	368	1.05	0.88	1.26	0.59
		TT	49	0.96	0.66	1.4	0.83
Lymphotoxin alpha	rs909253	CC	284	1.00	NA	NA	
		CT	685	0.90	0.74	1.1	0.32
		TT	422	1.02	0.81	1.28	0.85
Tumor Necrosis Factor	rs1799964	TT	896	1.00	NA	NA	
		CT	438	0.91	0.77	1.09	0.33

Gene name	rs number	Genotype	Number of participants	IRR	LCL	UCL	p_value
		GG	64	0.74	0.49	1.11	0.14
Tumor Necrois Factor	rs1800629	GG	1159	1.00	NA	NA	
		AG	225	1.24	1.00	1.53	<0.05
		AA	11	0.88	0.43	1.8	0.73
Tumor Necrosis Factor	rs361525	GG	1217	1.00	NA	NA	
		AG	171	1.08	0.84	1.4	0.54
		AA	10	0.49	0.29	0.84	<0.01
Tumor Necrosis Factor	rs3093662	AA	1068	1.00	NA	NA	
		AG	282	1.15	0.95	1.4	0.16
		GG	22	0.56	0.36	0.86	<0.01
Choline transporter-like protein 4	rs2242665	AA	784	1.00	NA	NA	
		AG	496	0.96	0.80	1.16	0.7
		GG	80	0.93	0.67	1.3	0.69
Interleukin 20 RA	rs1555498	CC	422	1.00	NA	NA	
		CT	690	0.92	0.76	1.11	0.4
		TT	284	1.00	0.79	1.26	0.99
NOD1	rs2075820	GG	449	1.00	NA	NA	
		AG	670	1.09	0.90	1.32	0.39
		AA	276	1.10	0.86	1.41	0.45
CD36	rs3211938	TT	1179	1.00	NA	NA	
		GT	210	0.78	0.61	1	0.05
		GG	8	0.15	0.02	1.23	0.08
CFTR	rs17140229	TT	687	1.00	NA	NA	
		AG	601	1.25	1.05	1.49	0.01
		CC	108	1.31	0.96	1.78	0.08
TLR4	rs4986790	AA	1169	1.00	NA	NA	
		AG	217	0.89	0.71	1.1	0.27
		GG	10	0.28	0.16	0.51	<0.0001
TLR4	rs4986791	CC	1379	1.00	NA	NA	

Gene name	rs number	Genotype	Number of participants	IRR	LCL	UCL	p_value
		CT	17	1.11	0.47	2.6	0.81
		TT	0	–	–	–	
TRIM5	rs7935564	GG	409	1.00	NA	NA	
		AG	710	0.83	0.69	0.99	0.04
		AA	192	1.16	0.88	1.54	0.3
RTN3	rs542998	TT	522	1.00	NA	NA	
		CT	650	1.01	0.84	1.22	0.93
		CC	223	1.20	0.92	1.56	0.17
IL22	rs2227507	TT	1313	1.00	NA	NA	
		CT	81	1.25	0.88	1.79	0.22
		CC	3	0.00	0.00	0	<0.0001
IL22	rs1012356	TT	352	1.00	NA	NA	
		AT	696	1.10	0.89	1.36	0.39
		AA	343	0.98	0.78	1.23	0.85
IL22	rs2227491	CC	501	1.00	NA	NA	
		CT	684	1.05	0.88	1.26	0.58
		TT	212	0.89	0.69	1.14	0.34
IL22	rs2227485	CC	504	1.00	NA	NA	
		AG	679	1.12	0.93	1.36	0.22
		AA	205	1.18	0.92	1.52	0.2
IL22	rs2227478	AA	220	1.00	NA	NA	
		AG	639	0.97	0.74	1.27	0.82
		GG	452	1.04	0.79	1.36	0.8
SPTB	rs229587	TT	610	1.00	NA	NA	
		CT	620	1.00	0.84	1.2	0.97
		CC	168	1.13	0.88	1.44	0.34
ADCY9	rs2230739	AA	1216	1.00	NA	NA	
		AG	175	0.89	0.68	1.17	0.41

Gene name	rs number	Genotype	Number of participants	IRR	LCL	UCL	p_value
		GG	6	3.59	1.88	6.82	<0.0001
ADCY9	rs10775349	CC	840	1.00	NA	NA	
		CG	486	0.98	0.82	1.18	0.87
		GG	70	1.02	0.73	1.4	0.93
IL4R	rs1805015	TT	399	1.00	NA	NA	
		CT	708	1.25	1.03	1.52	0.02
		CC	288	1.14	0.88	1.47	0.31
ADORA2B	rs2535611	TT	16	1.00	NA	NA	
		CT	281	0.58	0.33	1.02	0.06
		CC	1099	0.78	0.46	1.33	0.37
NOS2A	rs2297518	GG	1069	1.00	NA	NA	
		AG	310	0.86	0.70	1.06	0.15
		AA	19	1.28	0.75	2.19	0.37
NOS2A	rs9282799	CC	1255	1.00	NA	NA	
		CT	140	0.96	0.74	1.26	0.79
		TT	0	–	–	–	
NOS2A	rs8078340	CC	916	1.00	NA	NA	
		CT	432	0.99	0.83	1.19	0.96
		TT	49	1.58	1.01	2.46	0.04
EMR1	rs373533	GG	387	1.00	NA	NA	
		GT	715	0.92	0.75	1.12	0.4
		TT	293	0.77	0.61	0.97	0.02
ICAM1	rs5498	AA	1163	1.00	NA	NA	
		AG	218	1.00	0.79	1.25	0.97
		GG	10	0.99	0.46	2.11	0.97
GNAS	rs8386	CC	1007	1.00	NA	NA	
		CT	364	0.91	0.76	1.1	0.33
		TT	26	1.22	0.46	3.2	0.69

Gene name	rs number	Genotype	Number of participants	IRR	LCL	UCL	p_value
DERL3	rs3177244	AA	372	1.00	NA	NA	
		AG	691	0.98	0.80	1.2	0.85
		GG	332	1.03	0.82	1.3	0.8

Covariates controlled for are age, gender, education, SES, distance to Lwak, year and season.

5.3.2.3 Sex-linked genes: Effect of 4 SNPs on malaria incidence

SNPs from sex linked genes were analysed separately from autosomal gene SNPs. In this study, these SNPs were in two genes, the G6PD deficiency SNPs (amino-acid positions 376 and 202 mutations) and the CD40LG SNPs rs3092945 and rs17424229. Incidence rate ratios for G6PD deficiency mutation have been reported in Chapter 4 (see section 4.4.4). CD40LG analyses are reported here. The data were stratified by gender and analysed for the effect in the respective SNPs, rs3092945 and rs17424229.

The rs3092945 affects malaria incidence in boys but not in girls: Among boys, using the TT genotype as the reference, CC individuals had a higher incidence of malaria (IRR 1.39; 1.10-1.76; $p<0.01$). However, there was no difference in girls, where neither the CT individuals (IRR 0.99; 0.77-1.26; $p=0.91$) nor the CC individuals (IRR 0.96; 0.68-1.35; $p=0.81$) gave evidence of a difference with the TT reference group.

The rs17424229 does not affect malaria incidence in boys or girls: The TT genotype was used as the reference genotype. Among boys, there was no difference in incidence of malaria in the CC individuals (IRR 0.76; 0.54-1.06; $p=0.11$). Among girls, neither the CT individuals (IRR 0.93; 0.73-1.18; $p=0.53$) nor the CC individuals (IRR 0.96; 0.38-2.46; $p=0.93$) gave evidence of a difference with the TT reference group.

5.3.2.4 Effect of multivariate SNPs on malaria incidence

All genes with p-value less than 0.1 in the univariate analysis were used as predictors in a multivariate model, in addition to covariates. 3 genetic polymorphisms were found to be significantly related to malaria incidence.

i) EMR1 rs373533: Compared to the GG group, TT individuals were protected against clinical malaria (IRR 0.67; 0.53-0.84; p<0.001). The genotype frequencies were 28%, 51% and 21% for GG, GT and TT respectively. The role of EMR1 and comparison to other studies was described in section 5.3.2.2.4 above.

ii) TLR_6 rs5743809: Compared to the TT group, the rare CC genotype individuals were protected against clinical malaria (IRR 0.31; 0.16-0.6; p<0.001). The genotype frequencies were 82%, 17% and 1% for TT, CT and CC respectively. The role of TLR-6 and comparison with other studies was discussed in section 5.3.2.2.1 above.

iii) ADCY9 rs2230739: Compared to the AA group, rare GG genotype individuals had increased susceptibility to malaria (IRR 5.12; 2.23-11.75; p<0.001) incidence of clinical malaria episodes. The genotype frequencies were 87%, 12% and 1% for TT, CT and CC respectively. The role of ADCY9 has been described in section 5.3.2.2.4 above.

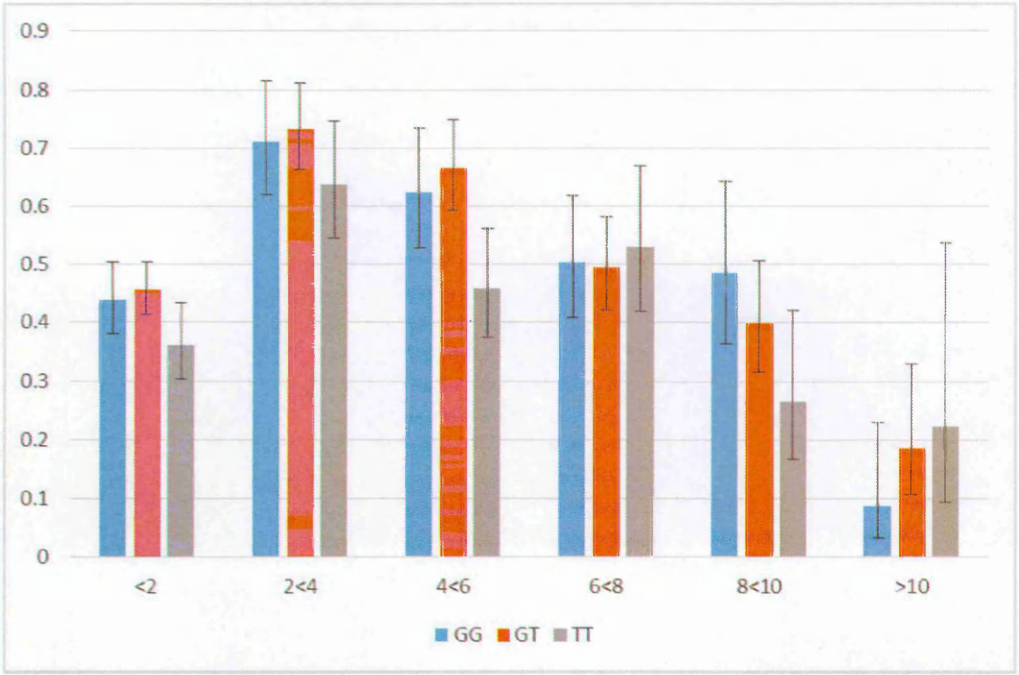
Among the genes above, the EMR1rs373533 variant had a protective variant (TT) with a high frequency. This suggests it is not deleterious, and may in fact be selected because of the protection it provides against malaria incidence. I therefore further investigated the age specific incidence of this gene in a Poisson-adjusted model, with and without covariates, as described in Table 5.4 below. Crude malaria incidence was also plotted according to genotype and age categories as shown in Figure 5.2

Table 5. 4 Age-specific incidence of malaria in EMR1 genotypes of rs373533

Age range	Genotypes	No. of participants(a)	No. of malaria episodes	cyfu	Crude malaria incidence(b)	Crude IRR (95% CI) #	P-value	Adjusted IRR (95% CI)	P-value
All	GG	365	681	1277.70	0.53	1		1	
	GT	673	1272	2312.20	0.55	1.03 (0.94-1.13)	0.50	0.93 (0.76-1.13)	0.45
	TT	278	462	1002.90	0.46	0.86 (0.77-0.97)	0.02	0.77 (0.62-0.97)	0.02
0-2	GG	298	191	435.58	0.44	1		1	
	GT	556	383	838.96	0.46	1.04 (0.88-1.24)	0.65	0.94 (0.71-1.24)	0.66
	TT	232	120	332.03	0.36	0.82 (0.66-1.04)	0.10	0.73 (0.51-1.03)	0.07
2-4	GG	176	207	291.23	0.71	1		1	
	GT	311	382	520.77	0.73	1.03 (0.87-1.22)	0.72	0.82 (0.62-1.09)	0.17
	TT	143	154	241.44	0.64	0.90 (0.73-1.11)	0.31	0.72 (0.50-1.02)	0.06
4-6	GG	149	141	226.50	0.62	1		1	
	GT	294	279	418.96	0.67	1.07 (0.87-1.31)	0.51	0.97 (0.70-1.35)	0.87
	TT	129	94	205.26	0.46	0.74 (0.57-0.96)	0.02	0.67 (0.46-0.98)	0.04
6-8	GG	114	90	179.03	0.50	1		1	
	GT	186	147	296.87	0.50	0.98 (0.76-1.28)	0.91	0.83 (0.56-1.23)	0.35
	TT	93	71	134.03	0.53	1.05 (0.77-1.44)	0.74	0.98 (0.63-1.52)	0.93
8-10	GG	67	48	99.09	0.48	1		1	
	GT	116	69	172.34	0.40	0.83 (0.57-1.19)	0.31	0.96 (0.54-1.72)	0.89
	TT	46	18	67.83	0.27	0.55 (0.32-0.94)	0.03	0.80 (0.40-1.60)	0.53
>=10	GG	35	4	46.31	0.09	1		1	
	GT	60	12	64.29	0.19	2.16 (0.70-6.70)	0.18	1.71 (0.58-5.05)	0.33
	TT	22	5	22.34	0.22	2.59 (0.70-9.65)	0.16	4.17 (1.06-16.35)	0.04

Adjusted rates refer to a multivariate model with the rs373533 and covariates age, gender, SES, mother's education, distance from hospital, year and season.

Figure 5. 2 Crude rates of uncomplicated malaria in EMR1 rs373533 genotypes stratified by age categories.



The y-axis represents the episodes of malaria per cyfu, whereas the x-axis represents the age-bands of the children in 2 year age categories. Error bars represent the 95% confidence interval for the estimate of malaria episodes per child per year. The blue bars represent children with EMR1 rs373533 GG genotype, orange bars represent children with EMR1 rs373533 GT genotype, while grey bars represent EMR1 rs373533 TT genotype.

Overall, the rs373533 TT variant reduced the incidence of uncomplicated malaria on both crude and adjusted analyses. The IRR for the adjusted analyses was 0.77 (95% C.I. 0.62-0.97; P<0.05). The other genes were not included in this analysis, hence the difference in the effect sizes seen between this model and the multi-gene model (see section 5.3.2.4). The current single gene analysis shows that the TT genotype was still significantly associated with a lower incidence of malaria, similar to the multigene model. Age specific investigation revealed significant protection at 4-6 years, but this effect appeared to be reversed for those over 10 years of age. However, the number of individuals over 10 years was low and therefore lacked sufficient power to conclude that the TT gene variant increases the incidence of malaria in this age group.

5.4 Discussion

This study investigated the protective effect that may arise from candidate malaria protective genes in a population of Lwak, Asembo. The major hypothesis under consideration was that variants of these malaria candidate genes significantly affect the incidence rate of malaria. I found that markers in 17 genes were related to malaria (see Table 5.2 above). However, some of these findings should be interpreted with caution, as the protective genotype frequency is <5%. Few studies consider susceptibility to uncomplicated malaria as a major outcome in malarial disease susceptibility studies. In contrast, this study contributes to the discourse on malaria susceptibility genes by focussing on susceptibility to mild malaria. One of the major results is that the same genes involved in susceptibility to severe malaria also implicated in susceptibility to mild malaria. These findings could be interpreted in at least two ways; first, that children at higher susceptibility to mild malaria are also much more likely to come down

with severe malaria. Alternatively, it may be that malaria progresses from asymptomatic to mild and eventually severe in a particular subset of children who carry a subset of genetic variants. These possibilities remain to be proved in mechanistic studies. However, there are important differences between this study and previous studies showing the risk of malaria in these gene variants. These previous studies were case-control studies in severe malaria susceptibility (Dunstan *et al.*, 2012; Kariuki *et al.*, 2013; Manjurano *et al.*, 2012), whereas this study was an observational cohort study whose main outcome was mild malaria.

The current study also considered potential sources of bias and confounding. Covariates of malaria incidence, a possible source of confounding, were identified through univariate analysis and adjusted for in the multivariate model. However, not all individuals had data on all the covariates, and therefore only the ones which had complete data for were included in the adjusted rates. Bias may arise in the conclusion of this study due to correlation of outcome measures from the same individual and relationships between the children studied. I used the sandwich estimator to account for possible bias due to the same individual contributing to multiple episodes. I sought to add an additional adjustment factor for household level clusters which approximate genetic relationships since I studied siblings, but the data showed that adjustment using the sandwich estimator was not significantly different from analyses using both the sandwich estimator and household cluster. The data were therefore only adjusted using the sandwich estimator.

The current approach used may also add value to genome wide association studies (GWAS). Although a candidate gene approach was used, several of the genes showing an appreciable

effect on univariate analysis were also analysed together in one Poisson regression model. However, the emphasis in GWAS studies is usually to identify new markers of resistance, and the complexity of GWAS means that complex statistical procedures are employed to choose the markers of significance (Jallow *et al.*, 2009). By comparing a few marker genes in a candidate gene approach, this study hopes to increase the chances of finding a true relation between the gene and malaria, un-confounded by the many SNP variations that may occur randomly in genome wide association comparisons.

Another important finding of the current study was that without controlling for effect of other genes, variants of thalassaemia and sickle both appear to significantly affect the incidence of uncomplicated malaria. However, when other candidate genes were controlled for, sickle and thalassaemia did not show a statistically significant effect on mild malaria. This may be due to interactions between the various genes. This possibility is considered in chapter 6.

It is clear from the list of significant genes that immuno-genetics plays a crucial role in protection from malaria. A review of malaria immunity was considered in order to set in context the findings in the current chapter. Immunity to malaria is complex; however, a few key features of malaria immunity are worth noting. Firstly, anti-malarial immunity is not sterile i.e. individuals may still fall ill to malaria when exposed (Marsh & Kinyanjui, 2006). Moreover, individuals who reside in malaria endemic areas appear to lose immunity when they travel to a non endemic site and subsequently fall ill more often when they return to their original endemic area. This suggests that constant exposure to the malaria parasite affords some immunity in the patients, and the phenomenon is termed as premunition (Smith

et al., 1999).

Secondly, anti-malarial immunity may further be divided into two components: anti-disease immunity, which is immunity to symptomatic effects of parasite infection and anti-parasite immunity, which is immunity that results in lower levels of parasitaemia (Marsh & Kinyanjui, 2006). Observations from endemic areas have shown that individuals develop tolerance to the malaria parasite, and therefore anti-disease immunity before the more specific anti-parasite immunity. Such anti-malarial immunity may be species and strain specific (Jeffery, 1966).

Finally, the mechanisms of immunity to malaria may explain premunition, anti-disease immunity, anti-parasite immunity and the observations in the current chapter. Malaria immune mechanisms have been broadly classified according to the components of the hosts' immunity or according to the parasite stages affected by the immune response. The host immune response involves both the cell mediated and humoral immune components. Cell mediated immunity is characterized by $CD8^+$ T cells, $CD4^+$ T cells, natural killer cells(NK), natural killer T cells (NKT), γδT cells and macrophages, whereas humoral immunity is mediated by antibodies produced by plasma cells and memory B cells. In a mouse model, within 48 hours of sporozoite inoculation, *Plasmodium* specific $CD8^+$ T cells are detectable in lymph nodes that drain the skin (Chakravarty *et al.*, 2007). Such T cells produce IFN-γ, which induces innate nitric oxide, a toxic radical which has been shown to limit the development or replication of malaria parasites (Schofield & Mueller, 2006).

In the liver, sporozoite inoculation studies have highlighted the importance of cell mediated immune (CMI) responses, where sterilizing immunity to infection by irradiated sporozoites has been achieved by inducing diverse CMI responses. The initial response is thought to result from innate recognition of the pathogen and reaction through natural killer and natural killer T cell mediated killing of infected cells. Adaptive responses result when the antigens from the parasites are processed and presented on MHC molecules, leading to $CD8^+$ T cells acting as cytotoxic T lymphocytes (CTL) by recognizing parasite infected cells and killing them. Dendritic cells may be crucial to this process, as they can act as an antigen presenting cells by phagocytizing sporozoite antigen and presenting peptides from parasites on MHC class I to $CD8^+$ T cells (Cockburn *et al.*, 2011). CD4+ T cells also produce IFN-γ, which promotes NK activity.

In the blood stages, the innate response is characterized by induction of macrophages, NK and NKT cells producing NO through the inducible nitric oxide synthase(iNOS), which are toxic to malaria infected cells (Schofield & Mueller, 2006). However, antibody responses are the most important at this stage in reducing parasitaemia and relieving symptoms of malaria, as shown in the classical immune antibody transfer from protected adults to sick children (Cohen *et al.*, 1961). Such antibodies are produced by plasma B cells and form the plasma antibodies that protect against malaria in endemic areas. However, such antibodies are known to decline rapidly in the absence of constant exposure (Weiss *et al.*, 2012). Memory B cells are also produced as a result of exposure to malaria parasites, and are generally more long lasting than circulating antibodies. In the absence of exposure, such memory B cells can induce rapid boosting of the antibody response when the host is re-exposed to malaria

antigens after several years (Wipasa *et al.*, 2010) and (Ndungu *et al.*, 2012). Repeated infections have also been implicated in the expansion of the memory B cell compartment of human immunity to malaria (Weiss *et al.*, 2012).

CD4+ T cells produce Th1 inflammatory cytokines, and later Th2 regulatory cytokines. Th1 pro-inflammatory cytokines and chemokines are associated with fever and other malarial symptoms (Collins & Jeffery, 1999). These include IFN-γ, TNF, IL-1β, IL-6, IL-12(p70) and IL-8 (Lyke *et al.*, 2004; Walther *et al.*, 2006). Inflammation results from interaction parasite derived products or by products interacting with immune cells, for example, haemozoin which binds to activating TLR9 (Parroche *et al.*, 2007) and GPI anchors to TLR2 (Krishnegowda *et al.*, 2005) trigger inflammation. Inflammation may also result from an interaction with host receptors, for example, some PfEMP1 variants bind endothelial protein C receptor in the brain(EPCR), thus blocking the protective function of the receptor in binding activated protein C and resulting in cerebral malaria (Turner *et al.*, 2013).

Since excessive inflammation can lead to severe and fatal malaria (Grau *et al.*, 1989), Th2 regulatory cytokines are also produced by the host. Work in mouse models has shown that regulatory cytokines TGF-β (Omer & Riley, 1998) and IL-10 (Li *et al.*, 1999) counter the production of pro-inflammatory cytokines –IFN γ and TNF. In humans, higher ratios of pro to anti-inflammatory cytokines in serum during acute malaria have been associated with more severe disease in some studies (Day, 1999; Dodoo *et al.*, 2002) but another study found a different result (Walther *et al.*, 2009). It has been argued that the key distinction of survival from life threatening malaria in early life may be linked to the ability to control excessive *P.*

falciparum–induced inflammation, since antibodies are only reliably acquired after many years of *P. falciparum* exposure (Portugal *et al.*, 2013).

The lack of sterilizing immunity in human malaria is likely to be due to many factors, which may include the high immuno-regulatory environment of the skin and the liver, (Honda *et al.*, 2011), the high levels of intra-hepatic response required to kill intra-hepatic parasites immunity and initiate sterilizing immunity (Schmidt *et al.*, 2010). In addition, studies show that the magnitude of vaccine-induced CD8+ specific for pre-erythrocytic antigens is correlated to protection (Webster *et al.*, 2005), and therefore sterilizing immunity is likely to be associated with responses above a high threshold which has not been achieved to date in natural infections.

Based on an understanding of the immune response in malaria, various vaccines have been formulated and studied for the protection they afford against malaria. The vaccine candidates target different stages of the life cycle of the parasites, as reviewed by Schwartz and colleagues (Schwartz *et al.*, 2012). Many of these target the blood stage as it produces symptoms of malaria, whereas some target the sporozoite based on the early promise of whole cell irradiated sporozoite induced protection (Nussenzweig *et al.*, 1967). The most advanced vaccine candidate to date is the circumsporozoite based *P. falciparum* RTS,S vaccine which confers sterile protection to than more half of malaria naïve adults (Casares *et al.*, 2010). In endemic populations, the vaccine reduces incidence of malaria by approximately 50% in children 5-17 months old (Agnandji *et al.*, 2011) and 30% in infants 6-12 weeks old (Agnandji *et al.*, 2012). Since the current study has shown genetic variants

that affect malaria incidence, evaluation of the impact of the vaccine may benefit from an analysis of the genetic variants of children in the various randomization arms.

5.5 Summary

This study found that 17 SNPs (see Table 5.2 above) made a significant difference in incidence rate ratio of malaria. In multivariate analysis, a polymorphism in EMRI-1 gave a significant result for protection from malaria. SNPs in two other genes, ADCY9 and TLR-6, had low numbers of protective or susceptibility loci. These findings should therefore be interpreted with caution since the power of the study was reduced by the low numbers of genotypes of variants affecting malaria incidence. The role of the genes has been identified from the literature, showing biological plausibility for why they could be related to malaria incidence. Interventions seeking to reduce malaria incidence, such as malaria vaccine studies, may benefit from an evaluation of the gene frequencies of some of these protective variants among study participants. Interactions between the various genes could be one reason why fewer genetic effects were seen in the multiple gene analysis. This possibility is investigated further in the next section, Chapter 6.

CHAPTER 6: Effect of interactions between malaria candidate genes on malaria incidence

6.1 Introduction

Background: Apart from the effect of individual genes, gene interaction is likely to be an important factor influencing malaria outcomes. Such genetic interaction, termed epistasis may play a role in complex trait outcomes (Carlborg & Haley, 2004), such as malaria. In recent years, examples of the effect of gene epistatic interactions on malaria outcomes have been shown in malaria endemic populations (Atkinson *et al.*, 2014; Williams, Mwangi, Wambua, Peto, *et al.*, 2005). These examples involve genes related to the red blood cell metabolism which interact to increase or reduce malaria severity. A brief overview of the role of epistasis in malaria incidence was presented in chapter 1 (section 1.2.4.5). In this chapter, I investigate two way genetic interactions, first between loci in red blood cell genes and second between loci in other malaria candidate genes in the Asembo genetic cohort.

Methods: I recruited a paediatric cohort of 1462 children from Asembo in western Kenya, and followed them up for incidence of malaria episodes by passive surveillance at Lwak mission hospital between 2008 and 2013. Detailed descriptions of the cohort, the study area and recruitment methods are outlined in Chapter 2. A multiplex mass array method was used

to type for all the genes tested for interaction, as described in section 2.6.1.3 and Appendix IV. In addition α^{+}thalassaemia was typed by PCR (as described in section 2.6.1.2) and added to this set for a final set of 68 genetic variant sets in 39 genes. I conducted analyses for the effect of interactions between variants of the typed genes as described below.

Statistical analysis: I fitted two models to genotype and malaria incidence data as follows:

1. A non-interaction Poisson model (Model I). This model predicts malaria incidence using, as dependent variables, a pair of genes and covariates such as age, distance to hospital, year, season, education, gender and socio-economic status.
2. An interaction Poisson model (Model II). It is similar to model I, but in addition has an interaction term representing the interaction between the pair of genes.

I used a likelihood ratio test (LRT) to compare the likelihood estimates of Model I and Model II. I inferred a significant interaction effect when the LRT p-value was < 0.05 or when the p-value was less than a Bonferroni correction of this p-value when multiple tests were performed.

Inferences: Models which showed a statistical significance were assessed for biological plausibility by checking biological databases of protein interactions and literature on the interacting pair of genes. Some hypotheses on why an interaction between the genes may affect malaria incidence were proposed.

6.2 Specific objectives

1. To investigate how co-inheritance of common red blood cell genes variants affect malaria incidence

2. To investigate the effect of co-inheritance at 68 malaria candidate gene loci on malaria incidence

Definitions:

1. **Statistical significance thresholds:** The threshold used to determine significantly interacting genes from the likelihood ratio test. Thresholds for objective 1 and 2, plus their justifications, are presented below.

Threshold considerations for objective 1: Significant interactions were assumed for those that had an LRT p-value < 0.05 since these interactions had been hypothesized a priori.

Threshold considerations for objective 2: Testing for multiple interactions is likely to introduce false positives. Therefore, a Bonferroni correction was used on the p-value by dividing it by the number of tests. In this comparison I did 56 tests on each of 57 SNPs, a total of 3192 tests. Therefore the significance p-value threshold was $< 0.05/3192 = 0.000016$

2. **Interaction terms:** A term denoting occurrence of two variants together, added to the Poisson model to check the effect of interaction. Examples of interaction terms are shown in table 6.1 for red blood cell genes. For example, for interaction between sickle and α thalassaemia, the interaction term 'a' in table 6.1 was used, whereas for interaction between sickle and CR1 Swain Langley, interaction term 'f' was used and so on.

Table 6. 1 Interaction terms for Sickle, ABO blood groups, G6PD deficiency in males, G6PD deficiency in females, CR1 McCoy blood groups and Swain Langley blood groups.

	α thalassaemia	ABO	G6PD def. male	G6PD def. female	CR1 McCoy	CR1 S. Langley
Sickle	Interaction term a	Interaction term b	Interaction term c	Interaction term d	Interaction term e	Interaction term f
Alpha thalassaemia		Interaction term g	Interaction term h	Interaction term i	Interaction term j	Interaction term k
ABO			Interaction term l	Interaction term m	Interaction term n	Interaction term o
G6PD deficiency male					Interaction term p	Interaction term q
G6PD deficiency female					Interaction term r	Interaction term s
McCoy						Interaction term t

def. = deficiency

6.3 Results

QC procedures. Since multiple testing using the LRT may give 'false positives' for interaction, stringent quality control procedures were applied on the resulting output. As outlined in chapter 5 (section 5.3.1), 11 SNPs were excluded from further analysis because they were either monomorphic (rs33930165, rs33950507, rs5743611, rs5743810, hcd36_g1439c and rs1799969), had minor allele frequencies less than 5% (rs2227507, rs9282799 and rs4986791) or showed a more significant deviation from Hardy-Weinberg equilibrium $p<0.001$ likely due to genotyping error (rs4986791 rs7935564 rs2814778).

To allow for biologically meaningful interpretations, RBC polymorphism interactions for which more is known are presented first (section 6.3.1). Once confidence in these results was

established, interactions between other genes in the same cohort were investigated (section 6.3.2) and possible promising pairs of interacting genes were proposed.

6.3.1 Red blood cell gene interactions

The results of interaction tests for red blood cell genes are reported in table 6.2 below.

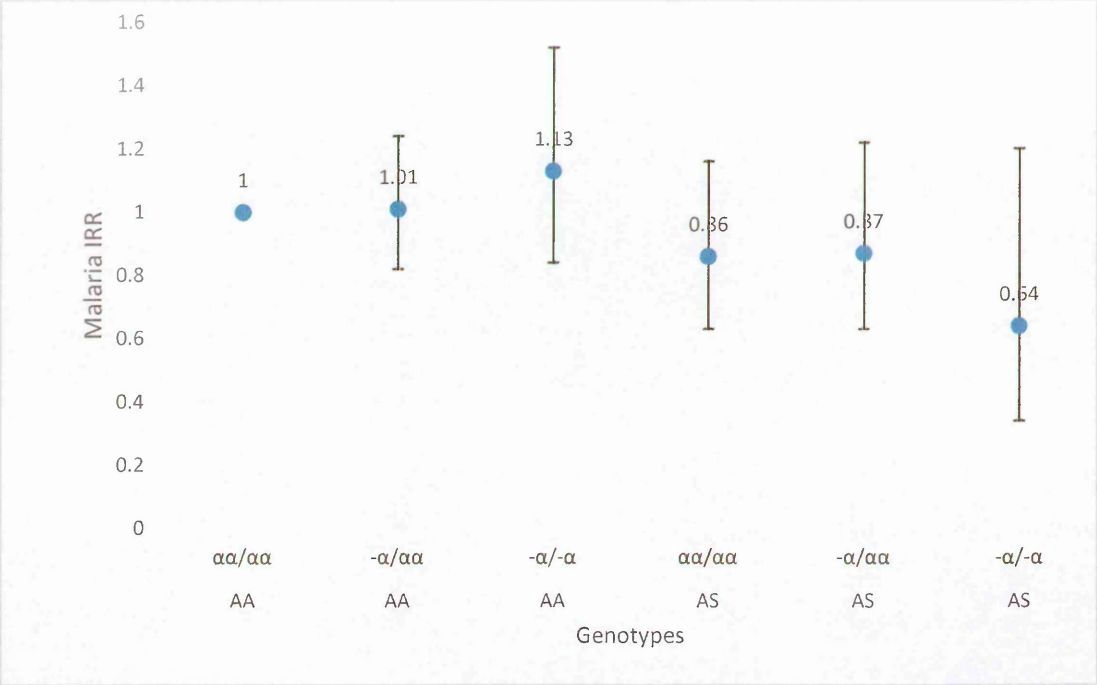
Table 6. 2 Likelihood ratio test p-values comparing a non-interaction models (model I) and a model with interaction (model II) for common genotypes of red blood cells.

	α thalassaemia	ABO	G6PD def. male	G6PD def. female	CR1 McCoy	CR1 S. Langley
Sickle	0.003	0.032	0.016	0.001	0.109	0.019
α-thalassaemia		<0.001	0.003	0.182	<0.001	0.008
ABO			0.015	<0.001	<0.001	<0.001
G6PD deficiency male					0.021	<0.001
G6PD deficiency female					0.086	0.031
McCoy						0.026

def. = deficiency

For the common red blood cell genes, there is evidence for interaction at the $p<0.05$ level. Results are presented to 3 decimal places to compare the relative evidence for interaction between the various gene pairs.

Figure 6. 1 Forest plot showing the interaction between sickle and thalassaemia



The effect of thalassaemia homozygotes in increasing the rate of malaria appears to be reversed in the background of sickle cell trait, such that homozygotes instead reduce the rate of malaria.

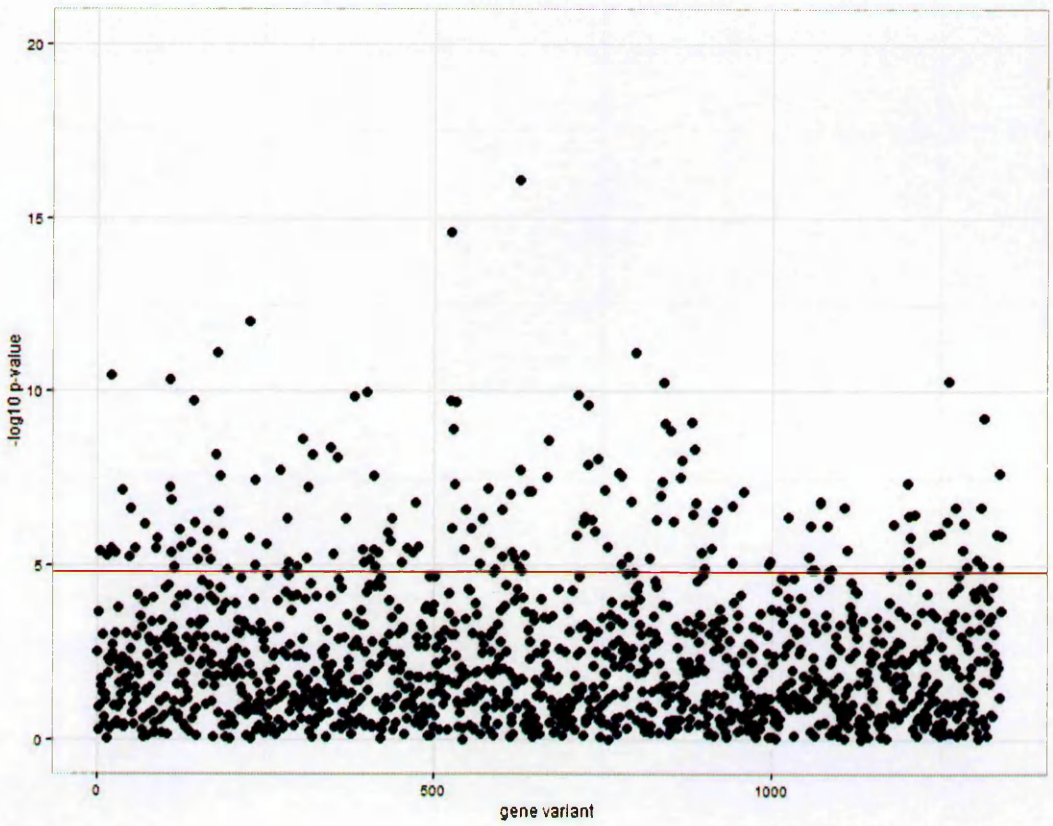
Table 6. 3 Incidence rate ratio (IRR), 95% lower class limit (LCL) and upper class limit (UCL) of various co-inheritance patterns of sickle and thalassaemia genotypes.

Sickle genotypes	Thalassaemia genotypes	IRR	LCL	UCL
AA	αα/αα	1		
	-α/αα	1.01	0.82	1.24
	-α/-α	1.13	0.84	1.52
AS	αα/αα	0.86	0.63	1.16
	-α/αα	0.87	0.63	1.22
	-α/-α	0.64	0.34	1.20

6.3.2 Interactions in 57 malaria candidate gene loci

Interaction tests were conducted between pairs of candidate gene loci, except the loci interacting with itself, resulting in a total of 57*56=3192 tests. The p-values for these tests were assessed if they were lower than the Bonferroni corrected value of 0.000016. Figure 6.2 is a Manhattan plot of all the p-values in the test, showing the ones above and below threshold.

Figure 6. 2 Manhattan plot of –log10 p-values for SNP interactions.



In the figure, values less than the threshold (0.000016 after Bonferroni correction) appear above the threshold value of 4.8 (the dashed line) since the p-values are transformed by multiplying by –log10.

Most of the p-values did not attain the threshold, but about 180 exceeded the threshold value, indicating a high incidence of potential statistical interaction between SNP pairs in modulating malaria incidence. The 50 lowest p-values are shown in table 6.4 below. They are arranged by the first gene (**gene 1)** of the interacting pair for ease of comprehending the number of gene interactions that one gene is likely to be involved in.

Table 6. 4 The top 50 SNP-pair hits detected using the LRT method. The table shows SNP pairs, SNP reference numbers and p-values of the LRT.

Gene 1	gene1_rsnumber	Gene2	gene2_rsnumber	LRT p-value
CR1	rs17047660	IL1B	rs1143634_2	2.43E-09
CR1	rs17047660	GNAS	rs8386	4.35E-09
CR1	rs17047660	TNF	rs1800629	6.66E-09
CR1	rs17047661	ADCY9	rs10775349_2	1.37E-10
CR1	rs17047661	C6	rs1801033_2	8.14E-09
IL1A	rs17411697	IL13	rs20541	1.05E-10
IL1A	rs17411697	NOD1	rs2075820_2	2.80E-08
TLR9	**rs352140**	**IL4**	**rs2243250**	**2.62E-15**
TLR9	rs352140	IL13	rs20541	1.92E-10
TLR9	rs352140	IL20RA	rs1555498_2	2.1E-10
TLR9	rs352140	TNF	rs1799964	1.26E-09
TLR9	rs352140	CTL4	rs2242665_2	4.57E-08
IL17RD	rs6780995	IL4	il4r63011_2	8.38E-17
IL17RD	rs6780995	NOS2	rs9282799_2	1.93E-08
TLR1	rs4833095	GNAS	rs8386	2.55E-09
TLR1	rs4833095	NOS2	hnos21659_2	3.00E-08
C6	rs1801033	TNF	rs1799964	1.35E-10
C6	rs1801033	IL22	rs2227491	2.5E-10
C6	rs1801033	IL22	rs1012356	1.3E-08
IRF1	rs2706384	CD40LG	rs3092945_2	4.39E-13
IRF1	rs2706384	IL13	rs20541	8.54E-09
IRF1	rs2706384	EMR1	rs373533_2	2.25E-08

Gene 1	gene1_rsnumber	Gene2	gene2_rsnumber	LRT p-value
IL13	rs20541	SPTB	rs229587_2	7.4E-12
IL13	**rs20541**	**IL4**	**rs2243250**	**2.60E-08**
IL4	rs2243250	DERL3	rs3177244_2	5.74E-11
IL4	rs2243250	NOS2	rs8078340	3.81E-08
LTA	rs2239704	LTA	rs909253	8.59E-10
LTA	rs2239704	CFTR	rs17140229_2	1.31E-09
LTA	rs2239704	ADORA2B	rs2535611_2	3.02E-08
LTA	rs909253	CFTR	rs17140229_2	7.46E-10
LTA	rs909253	RTN3	rs542998_2	4.62E-09
TNF	rs1800629	CD40LG	rs3092945_2	1.31E-09
NOD1	rs2075820	CD40LG	rs3092945_2	2.02E-08
RTN3	rs542998	DERL3	rs3177244_2	4.36E-08
IL22	rs2227491	CD40LG	rs3092945_2	3.69E-10
IL22	rs2227478	SPTB	rs229587_2	5.17E-11
ADORA2B	rs2535611	EMR1	rs373533_2	5.65E-10
ICAM1	rs5498	GNAS	rs8386	2.33E-08
GNAS	rs8386	CD40LG	rs3092945_2	1.55E-08
ABO	inferred_bg	RTN3	rs542998_2	8.55E-12
ABO	inferred_bg	IL4	rs1805015	8.53E-09
IL10	rs1800896	LTA	rs909253	1.01E-12
IL10	rs1800896	CFTR	rs17140229_2	3.49E-08
IL10	rs1800890	LTA	rs909253	1.89E-08
IL10	rs3024500	LTA	rs909253	7.79E-12
IL10	**rs3024500**	**IL4**	**rs2243250**	**6.63E-09**
IL10	rs3024500	TNF	rs361525	2.63E-08
GBP7	rs7537937	IL10	rs3024500	4.79E-11
GBP7	rs7537937	IL22	rs2227478	1.96E-10
HBB	rs334	LTA	rs909253	3.51E-11

Top 50 genetic interactions were those that had the lowest LRT p-value. Genes are arranged by the first gene of the interacting gene pair to show the various interactions one gene could have. SNPs highlighted in **bold** match the network of interactions identified using a biological database of interacting genes, the STRING database (see figure 6.3).

I expected a lower number of interaction signals. At least two reasons could account for the higher than expected significant p-values detected in this study (i) A high number of residual false positives (ii) Higher number of actual gene interactions than expected. These possibilities are considered in turn below.

(i) The analysis may reflect a high number of residual false positives

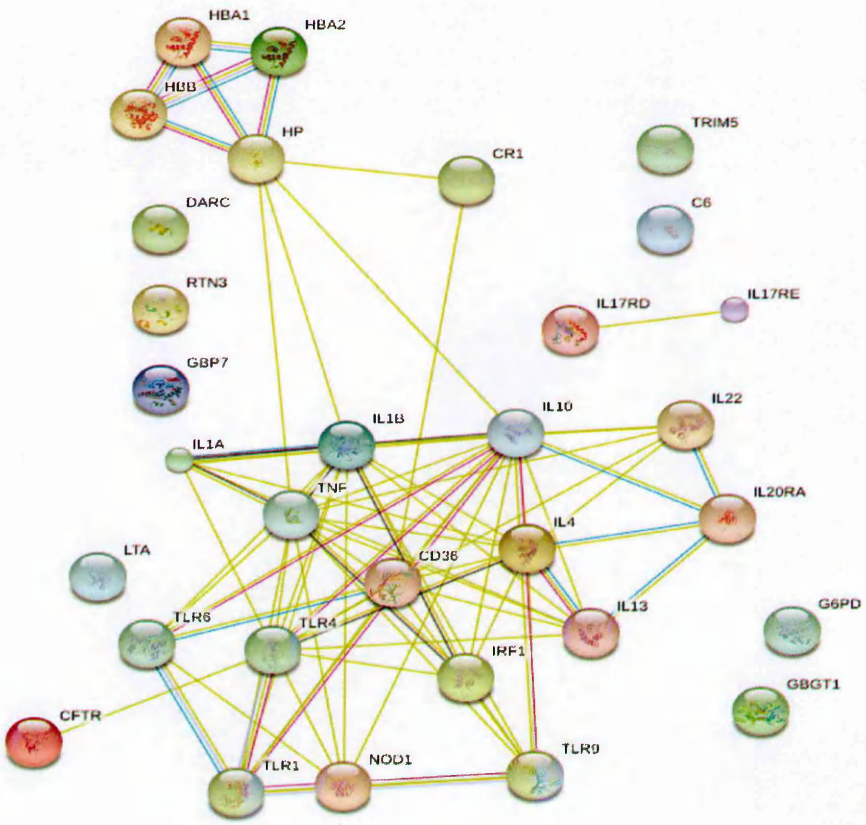
I used two approaches to address the possibility that the p-values detected were indeed false positives. The first was to compare the p-values detected with a known interaction which affects severe malaria incidence, the sickle cell trait and α-thalassaemia interaction. One important consideration is that the current study addresses uncomplicated malaria incidence. As such, the p-value for sickle and α-thalassaemia interaction as detected in this study for mild malaria was used as a comparator to any potentially significant interactions between genes that may affect mild malaria. The LRT p-value interaction was reported in section 6.3.1 above and yielded a p-value of 0.003 in this dataset (see section above). This was significant at the $p<0.05$ level. The p-values detected by the LRT were several orders of magnitude smaller than this p-value, ranging from 0.003 to 8.38E-17. It was possible that p-values much lower than 0.003 were spurious interaction effects. However, at the $p\leq0.003$ significance level, about 30% of the tests gave a signal for interaction, which I considered to be an unlikely possibility.

The second approach was to check if interactions between genes reported from this analysis had evidence from literature and biological databases. The tool used for this was the Search Tool for Retrieval of Interacting genes/proteins, STRING database, version 10. I submitted

31 protein names to the database for construction of a protein interaction network as seen in figure 6.3. The parameters used in the STRING database for the construction of the network were: medium confidence score of 0.4, evidence from all active prediction methods including neighbourhood, gene fusion, co-occurrence, co-expression, experiments, databases and text mining. Given its well documented role in malaria incidence, haptoglobin was included as part of the input to the STRING database to check for the proteins it interacts with, even though it was not typed in the current study. Interestingly, the STRING network demonstrates that haemoglobin related proteins (Haemoglobin beta (HBB), Haemoglobin alpha 1 (HBA1), Haemoglobin alpha2 (HBA2) and Haptoglobin (HP)) for which there is evidence of epistatic interaction for malaria incidence (Atkinson *et al.*, 2014; Williams, Mwangi, Wambua, Peto, *et al.*, 2005) cluster together and form a tight 1:1 interaction node of the network. This is consistent with the conclusion that the STRING network is an important tool in interpretation of statistical data and in validating hypotheses of interaction between genes (Emily *et al.*, 2009). The top 50 interactions from the LRT as reported in table 6.4 were compared with the nodes of the interaction network from the output of STRING (figure 6.3). From this comparison, 3 interactions from table 6.4 (in bold) matched interacting nodes from the STRING database. These gene interaction pairs are **TLR9-IL4, IL4-IL13** and **IL10-IL4.** The presence of these nodes appears to validate the results of some of the LRT analyses. However, the vast majority are not supported by evidence from the STRING database and may represent statistical but not biologically plausible interaction. Alternatively, the STRING database may yet be updated with evidence of these genetic interactions in future.

A high number of false positives could also be a feature of the Poisson model used. Lettre and colleagues (Lettre *et al.*, 2007) have shown that this model performs well in detecting effects of SNPs that are inherited in an additive or co-dominant fashion. However, it does not maximize power for those genes where recessive effects contribute to the disease being studied.

Figure 6. 3 Network of interacting genes from among the 31 being tested for interactions.



The figure was generated from the STRING database, version 10. Purple lines indicate evidence from experiments, blue lines indicate homology of sequences, green indicates evidence from text-mining, dark green indicate association by neighbourhood, Red indicates association by gene fusion whereas blue indicates association in gene and protein databases.

(ii) There may be a higher number of actual gene interactions than expected.

A review of epistasis has suggested that it may be more common than previously assumed (Carlborg & Haley, 2004). However, given that there are potentially a very large number of interactions between these genes, the current results should be interpreted with caution. The number of associations detected is dependent on the threshold of significance assumed, with higher thresholds resulting in higher number of signals. This study used the most conservative method of correction, which was expected to yield a very low threshold, and consequently, a few plausible epistasis candidate pairs. However, one challenge in the current analysis is that Bonferroni correction does not take account of linkages between genes. It assumes that the multiple tests are independent from each other, whereas one individual may carry several SNPs at the same time, or there could be linkage disequilibrium between SNPs and therefore one SNP could be related to several others. In addition, many SNPs would not remain significant on application of the most conservative methods such as the Bonferroni method. Among the alternative methods, the permutation approaches are currently favoured in choosing p-value cut-offs for multiple genetic tests, but they are complex and computationally intensive to implement. In addition, they assume independence of the multiple tests. An adaptation of the permutation approach for use in correlated genetic data has been described (Gao *et al.*, 2008) and has been used in various data sets which consider the same SNPs like this study (Dunstan *et al.*, 2012; Kariuki *et al.*, 2013; Manjurano *et al.*, 2012). Since the p-values cut off depends on the genetic SNP variation data, genetic structure may be a factor in choosing the correct cut-off. Given the complexity involved in such determinations, I considered using a threshold of significance similar to ones used for

previous studies in local populations. A review of the literature identified two studies (Kariuki *et al.*, 2013; Manjurano *et al.*, 2012) that investigated the association between these same SNPs and severe malaria, and in populations that are expected to share the same genetic structure as participants in the current study. One study, conducted in Tanzania (Manjurano *et al.*, 2012) used permutation methods to arrive at a cut-off p-value of $p<0.02$, whereas the other, conducted in Kilifi, used the method described by Gao and colleagues (Gao *et al.*, 2008) that approximates the permutation method to arrive at a cut-off p-value of $p<0.009$. The average of these (0.0145) was considered as a significance threshold. However, a major disadvantage of this approach was that the model specification used in these previous studies were different, as all the genes were modelled together in a multivariable model. In addition, in these previous studies, a LRT was not performed to check for interaction between loci. After all these considerations, I chose to use the Bonferroni correction since it is a conservative method and is expected to give the lowest number of false positives.

6.4 Discussion

Gene interactions are important in conditions of complex aetiology such as malaria. Epistasis could explain the absence of 'novel' genetic associations in studies, apart from classical loci such as sickle cell trait or α-thalassaemia as reported by Mackinnon and colleagues (Mackinnon *et al.*, 2005). In addition, genetic interactions may explain differing effects of genes on malaria outcomes in different populations. An example is the case of Interleukin 12 beta gene (IL12B) and lymphotoxin alpha (LTA) where different malaria phenotypes such as hyper-parasitaemia, severe malarial anaemia or cerebral malaria are often found to be associated with IL12B gene(Marquet *et al.*, 2008; Ong'echa *et al.*, 2011; Phawong *et al.*,

2010) while the association between parasitaemia with LTA is not consistent between studies (Barbier *et al.*, 2008; Clark *et al.*, 2009; Randall *et al.*, 2010).

A review of the effect of haptoglobin in various studies has also shown that the conflicting effects of haptoglobin on severe malaria reported in different studies could be explained by epistatic effects between the haptoglobin phenotypes and α thalassaemia(Atkinson *et al.*, 2014).

A study of genetic interactions may lead to hypothesis on mechanisms of action of the interacting genes as shown in studies conducted on the coast of Kenya (Atkinson *et al.*, 2014; Williams, Mwangi, Wambua, Peto, *et al.*, 2005). Thus, sickle and α-thalassaemia have been hypothesized to have contrasting mechanisms of action (Williams, Mwangi, Wambua, Peto, *et al.*, 2005). It has been shown that in populations where both of these conditions occur, the percentage of the HbS component of haemoglobin is reduced in individuals with higher α-thalassaemia deletions (Mouele *et al.*, 2000). The protective effect of HbS may therefore be lost due to having a lower proportion of the HbS component when co-inheritance with α-thalassaemia occurs. In populations with high frequency of both HbS and α-thalassaemia, interventions to reduce the effect of α-thalassaemia could potentially be used to maintain high HbS component among those who co-inherit both, thus reducing malaria incidence or severity.

The interaction between haptoglobin and α-thalassaemia (Atkinson *et al.*, 2014) has also led to proposals of various scenarios of its mechanism of action which could potentially be used

for interventions. Atkinson and colleagues consider different scenarios which affect the haptoglobin and α-thalassaemia genotypes. Co-inheritance of α-thalassaemia and Hp 2-2 genotype has been associated with higher incidence of severe malaria compared to coinheritance with Hp 2-1 or Hp 1-1. One of the mechanisms proposed is that Hp2-2 may be less able to quench the high oxidative stress associated with α-thalassaemia. Drugs which allow a better management of oxidative stress among α-thalassaemia patients could potentially improve severe malaria outcomes in those who co-inherit α-thalassaemia with Hp2-2.

The current study proposes that loci in cytokines and genes relevant to innate immunity may interact to affect malaria incidence. These are TLR9-IL4, IL4-IL13 and IL10-IL4 interactions. It is worth noting that each of the SNPs in these interacting gene pairs significantly affected the incidence of malaria on univariate analysis (see Table 5.2), with the exception of the SNP in interleukin 10 (rs302500) which gave marginal evidence (p=0.06) of a significant effect (see Table 5.3). This suggests that a useful strategy of identifying possible interacting gene pairs is to select SNPs which show evidence of affecting the malaria outcome and considering the effect of their interaction on malaria incidence. The role of TLR 9 was discussed in Chapter 5 section 5.3.2.2.1 while the role of IL-4 and IL-13 in malaria have been discussed in Chapter 5, section 5.3.2.2.2. I probed for the association between IL-4 and IL-13 and found that the genes are closely linked, display high sequence homology and have functional similarities (Zurawski & de Vries, 1994), further reinforcing the evidence for epistasis between the two genes in predicting uncomplicated malaria incidence. Thus, both IL-4 and IL-13 induce differentiation of naïve T cells to Th2 cells and affect the balance

between Th1 and Th2 cytokines. This suggests a possible intervention against malaria may be to balance cytokine responses as suggested by a previous study (Dodoo *et al.*, 2002).

IL-10, the other gene for which significant interaction was found, is an anti-inflammatory cytokine, and low levels have previously been associated with severe malaria among Ghanaian children (Kurtzhals *et al.*, 1998). Another study among Gambian children (Wilson *et al.*, 2005) found that children who had a haplotype including the G mutation were protected from severe malaria, but no specific polymorphism on its own was associated with severe malaria. The finding of particular haplotypes being protective is indicative of protection only being seen in a particular combination of IL-10 haplotypes or protection due to a polymorphism, not genotyped in that study but that may be in linkage disequilibrium with the IL-10 polymorphism. Since the G polymorphism was associated with protection from severe malaria but in this study gave evidence of susceptibility to mild malaria, it is consistent with the natural vaccine mechanism of protection from severe malaria. In the current study, IL-10 was found to interact with IL-4 to affect malaria incidence. Among the many roles of IL-4 is a reduction in Th1 cells and induction of Th2, while IL-10 also downregulates the expression of Th1 cells. It is possible that IL4 and IL10 interact positively to complement each other in reducing malaria incidence. The mechanisms underlying these interactions are yet to be investigated, however, it is possible that they affect processes in the development of malarial disease, as shown by a previous study which showed that SNPs in other immune-regulatory genes (TNF, IL12B and TLR4) interact epistatically to increase *P. falciparum* parasite growth in vitro (Basu *et al.*, 2012).

The current study also showed another significant interaction, the IL4-TLR9 interaction. Interactions with TLR-9 are potentially important, given its role in the immune recognition of malaria parasites. For example, it has been shown that a malaria vaccine candidate using the TLR-9 CpG ligands as an adjuvant promotes the acquisition of *P. falciparum* specific memory B cells in malaria naïve individuals (Crompton *et al.*, 2009) but not in semi-immune adults (Traore *et al.*, 2009). Since IL-4 suppresses responses to TLR-9 stimulation (Sriram *et al.*, 2014) it may be possible for co-inheritance of various genotypes of IL-4 and TLR-9 to significantly affect malaria incidence as suggested in the current study.

The above example demonstrates that interaction studies are useful in identifying key pathophysiological processes to target for interventions. Since malaria is a systemic disease which induces several immune reactions, deciding which reactions are useful and which ones are not is a complex problem. SNP interaction studies may clarify which interactions are beneficial and which ones are harmful to malaria outcomes. Beneficial interactions may be those that reduce malaria incidence, while harmful interactions may increase malaria incidence. As such, studies such as the current one can potentially provide a rationale for designing interventions to boost beneficial interactions while suppressing harmful ones.

6.4 Summary and conclusions

The LRT test was used to test how epistasis between gene markers (SNPs) affects malaria incidence. Results were interpreted by comparing the predicted interacting pairs with those in a database of known and predicted gene and protein interactions (STRING), and also comparing the interaction effect prediction strength to that of an expected interacting loci

pair, α-thalassaemia and sickle. Through such comparisons, 3 new possible interacting pairs of genes were identified; TLR9-IL4, IL4-IL13 and IL10-IL4. The importance of epistasis in understanding the roles of the genes and the giving clues to the pathogenesis of disease was considered in the discussion above.

In conclusion, further studies on the role of epistasis may be conducted, including more complex multiple gene epistasis models. The three pairs of candidate genes with epistatic interactions identified in this study may be studied in other genetic cohorts to confirm if they actually interact to increase or reduce malaria incidence. Validation of the several other candidate pairs of epistasis identified in the current study may also be undertaken in future studies.

CHAPTER 7: GENERAL DISCUSSION AND CONCLUSIONS

Through this thesis, I have used demographic surveillance and infectious disease surveillance platforms to study genetic protection from malaria incidence. The major themes investigated are the age and genotype-specific incidence of malaria, their relation to naturally acquired immunity and the effect of interactions between gene variants by multivariate and interaction models. I began by studying the malaria epidemiology in the Asembo population in chapter 3 by longitudinal and cross sectional approaches, defining a pragmatic definition of malaria in the surveillance platform and comparing this to the parasitological case definition established in the cross sectional survey.

As a first step in understanding the role of genetic polymorphisms in conditioning malaria incidence, I studied the best characterized of such genes-the red blood cell polymorphisms. While these genes have been known to affect malaria severity, their effect on mild malaria has received less attention, yet mild malaria may be the reservoir of potentially severe malaria in populations, especially in non-endemic populations (White, 2008). In contrast, in endemic populations, the major reservoir is asymptomatic (Githeko *et al.*, 1992). In the current push towards malaria elimination, populations are expected to transition from endemic, where the main infection reservoir is asymptomatic, to less endemic populations, where the uncomplicated malaria reservoir will be more significant. Anti-malarial use targeting symptomatic individuals to reduce the transmission reservoir of malaria (White, 2008) may

therefore be influenced by the role of genetics in differential susceptibility to mild malaria, as only individuals susceptible to mild malaria will be targeted. To achieve success, such programmes need to target all individuals positive for malaria parasites, regardless of symptoms as reported in recent mass screening and treatment campaigns (Tiono *et al.*, 2013).

While the distribution of common malaria protective genes has been modelled at increasingly better resolutions globally (Howes *et al.*, 2011; Howes *et al.*, 2012; Piel *et al.*, 2010), the current study makes several important contributions. First, it confirms the effect of these classical red blood cell genes on mild malaria incidence, which was subsequently used as a benchmark for evaluating the effects of other candidate genes on malaria incidence. Secondly, the age specific incidence patterns show that the effect significantly varies with age, and therefore interpretation of several studies should distinguish studies conducted among children of different ages, since what may be protective in one age group may not be protective in another age group (Taylor *et al.*, 2012). Thirdly, these studies contribute to debate on mechanisms of protection, such as the proposed natural vaccine hypothesis (Williams *et al.*, 1996) or the immune basis to the protection afforded by non-immune gene red blood cell polymorphisms (Williams, Mwangi, Roberts, *et al.*, 2005).

In Chapter 5, I further studied 34 more recently identified red blood cell metabolism, cytokine and immune effector genes for which there is evidence in the literature of an effect on malaria. From the literature, what is clear is that the effects of candidate malaria genetic variants are not universal, what may be protective in one population may lead to susceptibility in another as shown by the case of LTA (Barbier *et al.*, 2008; Clark *et al.*, 2009; Randall *et*

al., 2010). Such differences may arise because of differences in transmission intensity, different background genetics or demographic factors. Thus, a study on genetic factors affecting malaria incidence in the Asembo population may be used to compare with effects reported in other settings. Further, I modelled them in a multivariate model to come up with a model that explains the best predictors of malaria in this population. Compared to the univariate models for each gene separately, the multivariate model is a better model of reality as it allows the variants to control for each other's effects while at the same time controlling for confounding due to covariates such as season, year, socio-economic status and attrition of the participants seen in the hospital due to distance from the hospital.

Incidence data in the various genotypes can be used to predict possible mechanisms of protection from severe malaria. Involvement of the same SNP in mild malaria in this study compared to severe malaria in other studies was interpreted as meaning the SNPs allow for a progression from mild to severe malaria. This could mean that at least the same processes leading to mild malaria, given time could lead to severe malaria. However, scenarios where one SNP enhances to susceptibility to mild malaria but protects from severe malaria as shown by the literature were encountered. This was interpreted as either contrasting mechanisms of protection in mild and severe malaria, or protection from severe malaria coming at a cost of inducing susceptibility to mild malaria. These mechanisms need not be mutually exclusive, as malaria may both be a disease that manifests progressively from mild to severe, and at the same time be a disease with alternative pathways of illness depending on the host i.e. selecting those patients who once infected may manifest either severe or mild disease symptoms. SNP markers associated with discernible changes in malaria risk also give

insights into the pathogenesis of malaria. Most of these markers were in immune mediator genes, suggesting the crucial role that these genes play in distinguishing between those who develop symptoms and those who are asymptomatic. In the current study, toll like receptors 4, 6 and 9 polymorphisms were associated with either increased or reduced risk of uncomplicated malaria. This highlights the relative importance of these genes in the innate response to malaria infection. Likewise, SNP markers in complement 6, IL-4, IL-13 which were related to mild malaria incidence indicate the role of immune effector mechanisms in determining who displays symptoms of malaria.

Among the reasons for difference in effects may be gene interactions. I investigated this possibility in chapter 6 by comparing interaction and non-interaction models to find if there was a significant difference. Such significant differences between the two models were inferred as a possible interaction between the respective genes in conditioning malaria susceptibility. Predictions of interacting alleles in this dataset were further confirmed using a literature database of genetic and protein interactions. Through this, I came up with 3 pairs of genes that have evidence both statistically in my dataset and external validation from biological databases and literature. These three pairs of genes represent candidates for epistatic interaction studies and may be further explored to show their effects in larger studies.

There were a number of limitations in this study, including a lack of quantitative parasite density data at the hospital. During the course of the study, malaria slides were collected mainly for qualitative determination of malaria status, as such parasite density data were not

available. A possible way of meeting this challenge going forward would be to collect DBS from each patient while making the malaria slides. DBS are easier to store than malaria slides and can allow for extraction of parasite DNA later, thus allowing for determination of parasite density. Nevertheless, several studies on malaria have used presence of parasites during febrile episodes and no other reason detected for hospital visits as a case definition for malaria. This definition is the same as the one normally used for clinical management of malaria. Thus, protective effects can be directly interpreted as representing a lower burden on health facilities, while susceptibility inducing effects show the excess burden that carriers of these polymorphisms present to the health facilities.

Another limitation was the passive follow up at the health facility as opposed to what would have been more ideal-active follow-up and diagnosis in the study participant's homes. Retrospective and prospective approaches were used in the current study. The bias this may cause may be due to inconsistency in collection of data between both approaches. However, I confirmed that the data collection methods were similar during the period, and there were no major changes that may impact the current study. However, temporal changes in transmission which may occur naturally or through interventions which may have been introduced in particular years was accounted for by adjusting for year and season in the model.

Recruitment of siblings may have meant the population was more related than would be found in a random sample. However, this data showed a negligible effect of adjustments for both clustering by household and then by individual episodes of data in the same individual

(the sandwich estimator). The results reported here are therefore those for the sandwich estimator alone.

Finally, this work forms a baseline for further studies, which may include:

i) Finding the optimal mix of effects which is necessary to provide malaria protection, or allow for the design of efficient vaccines. This may include haplotype analysis.

ii) Testing the interaction candidates identified in this study in larger studies powered to detect the interaction effect.

iii) Confirming the effects of the SNPs using rigorously defined severe malaria cases in this population.

iv) More complex epistatic gene interaction analyses, such as 3 way interactions.

v) Comparison of effects of malaria interventions in children of different genotypes.

vi) Since most of the SNPs involve the innate immune responses, they may not be specific to malaria and are likely to have an effect on other diseases. Future studies may be designed to test for genetic susceptibility to other diseases detected in the area.

REFERENCES

Abu-Zeid, Y. A., Abdulhadi, N. H., Theander, T. G., Hviid, L., Saeed, B. O., Jepsen, S., . . . Bayoumi, R. A. (1992). Seasonal changes in cell mediated immune responses to soluble Plasmodium falciparum antigens in children with haemoglobin AA and haemoglobin AS. *Trans R Soc Trop Med Hyg, 86*(1), 20-22.

Achidi, E. A., Salimonu, L. S., Perlmann, H., Perlmann, P., Berzins, K., & Williams, A. I. (1996). Lack of association between levels of transplacentally acquired Plasmodium falciparum-specific antibodies and age of onset of clinical malaria in infants in a malaria endemic area of Nigeria. *Acta Trop, 61*(4), 315-326.

Adazu, K., Lindblade, K. A., Rosen, D. H., Odhiambo, F., Ofware, P., Kwach, J., . . . Slutsker, L. (2005). Health and demographic surveillance in rural western Kenya: a platform for evaluating interventions to reduce morbidity and mortality from infectious diseases. *Am J Trop Med Hyg, 73*(6), 1151-1158.

Agnandji, S. T., Lell, B., Fernandes, J. F., Abossolo, B. P., Methogo, B. G., Kabwende, A. L., . . . Vansadia, P. (2012). A phase 3 trial of RTS,S/AS01 malaria vaccine in African infants. *N Engl J Med, 367*(24), 2284-2295.

Agnandji, S. T., Lell, B., Soulanoudjingar, S. S., Fernandes, J. F., Abossolo, B. P., Conzelmann, C., . . . Vansadia, P. (2011). First results of phase 3 trial of RTS,S/AS01 malaria vaccine in African children. *N Engl J Med, 365*(20), 1863-1875.

Aidoo, M., Terlouw, D. J., Kolczak, M. S., McElroy, P. D., ter Kuile, F. O., Kariuki, S., . . . Udhayakumar, V. (2002). Protective effects of the sickle cell gene against malaria morbidity and mortality. *Lancet, 359*(9314), 1311-1312. doi: 10.1016/s0140-6736(02)08273-9

Aikawa, M., Miller, L. H., Rabbege, J. R., & Epstein, N. (1981). Freeze-fracture study on the erythrocyte membrane during malarial parasite invasion. *J Cell Biol, 91*(1), 55-62.

Alcais, A., Quintana-Murci, L., Thaler, D. S., Schurr, E., Abel, L., & Casanova, J. L. (2010). Life-threatening infectious diseases of childhood: single-gene inborn errors of immunity? *Ann N Y Acad Sci, 1214*, 18-33.

Allen, S. J., Rowe, P., Allsopp, C. E., Riley, E. M., Jakobsen, P. H., Hill, A. V., & Greenwood, B. M. (1993). A prospective study of the influence of alpha thalassaemia on morbidity from malaria and immune responses to defined Plasmodium falciparum antigens in Gambian children. *Trans R Soc Trop Med Hyg, 87*(3), 282-285.

Allison, A. C. (1954). Protection afforded by sickle-cell trait against subtertian malareal infection. *Br Med J, 1*(4857), 290-294.

Allison, A. C., & Clyde, D. F. (1961). Malaria in African children with deficient erythrocyte glucose-6-phosphate dehydrogenase. *Br Med J, 1*(5236), 1346-1349.

Amek, N., Bayoh, N., Hamel, M., Lindblade, K. A., Gimnig, J. E., Odhiambo, F., . . . Vounatsou, P. (2012). Spatial and temporal dynamics of malaria transmission in rural Western Kenya. *Parasit Vectors, 5*, 86.

Amin, A. A., Zurovac, D., Kangwana, B. B., Greenfield, J., Otieno, D. N., Akhwale, W. S., & Snow, R. W. (2007). The challenges of changing national malaria drug policy to artemisinin-based combinations in Kenya. *Malar J, 6*, 72. doi: 10.1186/1475-2875-6-72

Anstey, N. M., Weinberg, J. B., Hassanali, M. Y., Mwaikambo, E. D., Manyenga, D., Misukonis, M. A., . . . Granger, D. L. (1996). Nitric oxide in Tanzanian children with malaria: inverse relationship between malaria severity and nitric oxide production/nitric oxide synthase type 2 expression. *J Exp Med, 184*(2), 557-567.

Apinjoh, T. O., Anchang-Kimbi, J. K., Njua-Yafi, C., Ngwai, A. N., Mugri, R. N., Clark, T. G., . . . Achidi, E. A. (2014). Association of candidate gene polymorphisms and TGF-beta/IL-10 levels with malaria in three regions of Cameroon: a case-control study. *Malar J, 13*, 236. doi: 10.1186/1475-2875-13-236

Aponte, J. J., Menendez, C., Schellenberg, D., Kahigwa, E., Mshinda, H., Vountasou, P., . . . Alonso, P. L. (2007). Age interactions in the development of naturally acquired immunity to Plasmodium falciparum and its clinical presentation. *PLoS Med, 4*(7), e242. doi: 10.1371/journal.pmed.0040242

Armitage P, B. G., Mathews JNS. (2001). *Statistical methods in medical research* (4th ed.). Oxford: Blackwell Scientific Publications.

Atkinson, S. H., Uyoga, S. M., Nyatichi, E., Macharia, A. W., Nyutu, G., Ndila, C., . . . Williams, T. N. (2014). Epistasis between the haptoglobin common variant and alpha+thalassemia influences risk of severe malaria in Kenyan children. *Blood, 123*(13), 2008-2016.

Auburn, S., Fry, A. E., Clark, T. G., Campino, S., Diakite, M., Green, A., . . . Kwiatkowski, D. P. (2010). Further evidence supporting a role for gs signal transduction in severe malaria pathogenesis. *PLoS One, 5*(4), e10017. doi: 10.1371/journal.pone.0010017

Aydin-Schmidt, B., Mubi, M., Morris, U., Petzold, M., Ngasala, B. E., Premji, Z., . . . Martensson, A. (2013). Usefulness of Plasmodium falciparum-specific rapid diagnostic tests for assessment of parasite clearance and detection of recurrent infections after artemisinin-based combination therapy. *Malar J, 12*, 349.

Ayi, K., Turrini, F., Piga, A., & Arese, P. (2004). Enhanced phagocytosis of ring-parasitized mutant erythrocytes: a common mechanism that may explain protection against falciparum malaria in sickle trait and beta-thalassemia trait. *Blood, 104*(10), 3364-3371. doi: 10.1182/blood-2003-11-3820

Baird, J. K. (1995). Host age as a determinant of naturally acquired immunity to Plasmodium falciparum. *Parasitol Today, 11*(3), 105-111.

Baird, J. K. (2009). Severe and fatal vivax malaria challenges 'benign tertian malaria' dogma. *Ann Trop Paediatr, 29*(4), 251-252.

Barbier, M., Delahaye, N. F., Fumoux, F., & Rihet, P. (2008). Family-based association of a low producing lymphotoxin-alpha allele with reduced Plasmodium falciparum parasitemia. *Microbes Infect, 10*(6), 673-679. doi: 10.1016/j.micinf.2008.03.001

Basu, M., Das, T., Ghosh, A., Majumder, S., Maji, A. K., Kanjilal, S. D., . . . Sengupta, S. (2012). Gene-gene interaction and functional impact of polymorphisms on innate immune genes in controlling Plasmodium falciparum blood infection level. *PLoS One, 7*(10), e46441.

Baum, J., Chen, L., Healer, J., Lopaticki, S., Boyle, M., Triglia, T., . . . Cowman, A. F. (2009). Reticulocyte-binding protein homologue 5 - an essential adhesin involved in invasion of human erythrocytes by Plasmodium falciparum. *Int J Parasitol, 39*(3), 371-380.

Bayoh, M. N., Mathias, D. K., Odiere, M. R., Mutuku, F. M., Kamau, L., Gimnig, J. E., . . . Walker, E. D. (2010). Anopheles gambiae: historical population decline associated with regional distribution of insecticide-treated bed nets in western Nyanza Province, Kenya. *Malar J, 9*, 62. doi: 10.1186/1475-2875-9-62

Bayoh, M. N., Walker, E. D., Kosgei, J., Ombok, M., Olang, G. B., Githeko, A. K., . . . Gimnig, J. E. (2014). Persistently high estimates of late night, indoor exposure to malaria vectors despite high coverage of insecticide treated nets. *Parasit Vectors, 7*, 380. doi: 10.1186/1756-3305-7-380

Bayoumi, R. A. (1987). The sickle-cell trait modifies the intensity and specificity of the immune response against P. falciparum malaria and leads to acquired protective immunity. *Med Hypotheses, 22*(3), 287-298.

Beet, E. A. (1946). Sickle cell disease in the Balovale District of Northern Rhodesia. *East Afr Med J, 23*, 75-86.

Bienzle, U., Ayeni, O., Lucas, A. O., & Luzzatto, L. (1972). Glucose-6-phosphate dehydrogenase and malaria. Greater resistance of females heterozygous for enzyme deficiency and of males with non-deficient variant. *Lancet, 1*(7742), 107-110.

Bigogo, G., Audi, A., Aura, B., Aol, G., Breiman, R. F., & Feikin, D. R. (2010). Health-seeking patterns among participants of population-based morbidity surveillance in rural western Kenya: implications for calculating disease rates. *Int J Infect Dis, 14*(11), e967-973. doi: 10.1016/j.ijid.2010.05.016

Bloland, P. B., Boriga, D. A., Ruebush, T. K., McCormick, J. B., Roberts, J. M., Oloo, A. J., . . . Campbell, C. C. (1999). Longitudinal cohort study of the epidemiology of malaria infections in an area of intense malaria transmission II. Descriptive epidemiology of malaria infection and disease among children. *Am J Trop Med Hyg, 60*(4), 641-648.

Bruce-Chwatt, L. J. (1981). Alphonse Laveran's discovery 100 years ago and today's global fight against malaria. *J R Soc Med, 74*(7), 531-536.

Burkot, T. R., & Graves, P. M. (1995). The value of vector-based estimates of malaria transmission. *Ann Trop Med Parasitol, 89*(2), 125-134.

Butcher, G. A., Garland, T., Ajdukiewicz, A. B., & Clark, I. A. (1990). Serum tumor necrosis factor associated with malaria in patients in the Solomon Islands. *Trans R Soc Trop Med Hyg, 84*(5), 658-661.

Butler, A. R., Khan, S., & Ferguson, E. (2010). A brief history of malaria chemotherapy. *J R Coll Physicians Edinb, 40*(2), 172-177.

Butler, T. (1973). G-6-PD deficiency and malaria in Black Americans in Vietnam. *Mil Med, 138*(3), 153-155.

Cappadoro, M., Giribaldi, G., O'Brien, E., Turrini, F., Mannu, F., Ulliers, D., . . . Arese, P. (1998). Early phagocytosis of glucose-6-phosphate dehydrogenase (G6PD)-deficient erythrocytes parasitized by Plasmodium falciparum may explain malaria protection in G6PD deficiency. *Blood, 92*(7), 2527-2534.

Carlborg, O., & Haley, C. S. (2004). Epistasis: too often neglected in complex trait studies? *Nat Rev Genet, 5*(8), 618-625.

Carpenter, D., Rooth, I., Farnert, A., Abushama, H., Quinnell, R. J., & Shaw, M. A. (2009). Genetics of susceptibility to malaria related phenotypes. *Infect Genet Evol, 9*(1), 97-103. doi: 10.1016/j.meegid.2008.10.008

Carter, N., Pamba, A., Duparc, S., & Waitumbi, J. N. (2011). Frequency of glucose-6-phosphate dehydrogenase deficiency in malaria patients from six African countries enrolled in two randomized anti-malarial clinical trials. *Malar J, 10*, 241. doi: 10.1186/1475-2875-10-241

Casares, S., Brumeanu, T. D., & Richie, T. L. (2010). The RTS,S malaria vaccine. *Vaccine, 28*(31), 4880-4894.

Ceesay, S. J., Casals-Pascual, C., Erskine, J., Anya, S. E., Duah, N. O., Fulford, A. J., . . . Conway, D. J. (2008). Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis. *Lancet, 372*(9649), 1545-1554.

Chakravarty, S., Cockburn, I. A., Kuk, S., Overstreet, M. G., Sacci, J. B., & Zavala, F. (2007). CD8+ T lymphocytes protective against malaria liver stages are primed in skin-draining lymph nodes. *Nat Med, 13*(9), 1035-1041.

Chitnis, C. E., & Sharma, A. (2008). Targeting the Plasmodium vivax Duffy-binding protein. *Trends Parasitol, 24*(1), 29-34.

Chong, S. S., Boehm, C. D., Higgs, D. R., & Cutting, G. R. (2000). Single-tube multiplex-PCR screen for common deletional determinants of alpha-thalassemia. *Blood, 95*(1), 360-362.

Clark, T. G., Diakite, M., Auburn, S., Campino, S., Fry, A. E., Green, A., . . . Kwiatkowski, D. P. (2009). Tumor necrosis factor and lymphotoxin-alpha polymorphisms and severe malaria in African populations. *J Infect Dis, 199*(4), 569-575. doi: 10.1086/596320

Cockburn, I. A., Tse, S. W., Radtke, A. J., Srinivasan, P., Chen, Y. C., Sinnis, P., & Zavala, F. (2011). Dendritic cells and hepatocytes use distinct pathways to process protective antigen from plasmodium in vivo. *PLoS Pathog, 7*(3), e1001318.

Cohen, S., Mc, G. I., & Carrington, S. (1961). Gamma-globulin and acquired immunity to human malaria. *Nature, 192*, 733-737.

Coleman, B. I., Skillman, K. M., Jiang, R. H., Childs, L. M., Altenhofen, L. M., Ganter, M., . . . Duraisingh, M. T. (2014). A Plasmodium falciparum histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host Microbe, 16*(2), 177-186.

Collins, W. E., & Jeffery, G. M. (1999). A retrospective examination of secondary sporozoite- and trophozoite-induced infections with Plasmodium falciparum: development of parasitologic and clinical immunity following secondary infection. *Am J Trop Med Hyg, 61*(1 Suppl), 20-35.

Courtin, D., Milet, J., Bertin, G., Vafa, M., Sarr, J. B., Watier, L., . . . Migot-Nabias, F. (2011). G6PD A-variant influences the antibody responses to Plasmodium falciparum MSP2. *Infect Genet Evol, 11*(6), 1287-1292. doi: 10.1016/j.meegid.2011.04.016

Crompton, P. D., Mircetic, M., Weiss, G., Baughman, A., Huang, C. Y., Topham, D. J., . . . Pierce, S. K. (2009). The TLR9 ligand CpG promotes the acquisition of Plasmodium falciparum-specific memory B cells in malaria-naive individuals. *J Immunol, 182*(5), 3318-3326. doi: 10.4049/jimmunol.0803596

D'Acremont, V., Lengeler, C., & Genton, B. (2010). Reduction in the proportion of fevers associated with Plasmodium falciparum parasitaemia in Africa: a systematic review. *Malar J, 9*, 240. doi: 10.1186/1475-2875-9-240

Day, N. P. (1999). The prognostic and pathophysiologic role of pro- and antiinflammatory cytokines in severe malaria. *J Infect Dis, 180*(4), 1288-1297.

Desai, M., Buff, A. M., Khagayi, S., Byass, P., Amek, N., van Eijk, A., . . . Hamel, M. J. (2014). Age-specific malaria mortality rates in the KEMRI/CDC health and demographic surveillance system in western Kenya, 2003-2010. *PLoS One, 9*(9), e106197. doi: 10.1371/journal.pone.0106197

Dodoo, D., Omer, F. M., Todd, J., Akanmori, B. D., Koram, K. A., & Riley, E. M. (2002). Absolute levels and ratios of proinflammatory and anti-inflammatory cytokine production in vitro predict clinical immunity to Plasmodium falciparum malaria. *J Infect Dis, 185*(7), 971-979. doi: 10.1086/339408

Dunstan, S. J., Rockett, K. A., Quyen, N. T., Teo, Y. Y., Thai, C. Q., Hang, N. T., . . . Hien, T. T. (2012). Variation in human genes encoding adhesion and proinflammatory molecules are associated with severe malaria in the Vietnamese. *Genes Immun, 13*(6), 503-508. doi: 10.1038/gene.2012.25

Dvorak, J. A., Miller, L. H., Whitehouse, W. C., & Shiroishi, T. (1975). Invasion of erythrocytes by malaria merozoites. *Science, 187*(4178), 748-750.

Emily, M., Mailund, T., Hein, J., Schauser, L., & Schierup, M. H. (2009). Using biological networks to search for interacting loci in genome-wide association studies. *Eur J Hum Genet, 17*(10), 1231-1240. doi: 10.1038/ejhg.2009.15

Eng, B., Ainsworth, P., & Waye, J. S. (1994). Anomalous migration of PCR products using nondenaturing polyacrylamide gel electrophoresis: the amelogenin sex-typing system. *J Forensic Sci, 39*(6), 1356-1359.

Esposito, S., Molteni, C. G., Zampiero, A., Baggi, E., Lavizzari, A., Semino, M., . . . Principi, N. (2012). Role of polymorphisms of toll-like receptor (TLR) 4, TLR9, toll-interleukin 1 receptor domain containing adaptor protein (TIRAP) and FCGR2A genes in malaria susceptibility and severity in Burundian children. *Malar J, 11*, 196. doi: 10.1186/1475-2875-11-196

Eziefula, A. C., Gosling, R., Hwang, J., Hsiang, M. S., Bousema, T., von Seidlein, L., & Drakeley, C. (2012). Rationale for short course primaquine in Africa to interrupt malaria transmission. *Malar J, 11*, 360.

Feachem, R. G., Phillips, A. A., Hwang, J., Cotter, C., Wielgosz, B., Greenwood, B. M., . . . Snow, R. W. (2010). Shrinking the malaria map: progress and prospects. *Lancet, 376*(9752), 1566-1578.

Feikin, D. R., Olack, B., Bigogo, G. M., Audi, A., Cosmas, L., Aura, B., . . . Breiman, R. F. (2011). The burden of common infectious disease syndromes at the clinic and household level from population-based surveillance in rural and urban Kenya. *PLoS One, 6*(1), e16085. doi: 10.1371/journal.pone.0016085 [doi]

Frevert, U., Engelmann, S., Zougbede, S., Stange, J., Ng, B., Matuschewski, K., . . . Yee, H. (2005). Intravital observation of Plasmodium berghei sporozoite infection of the liver. *PLoS Biol, 3*(6), e192.

Friedman, M. J. (1978). Erythrocytic mechanism of sickle cell resistance to malaria. *Proc Natl Acad Sci U S A, 75*(4), 1994-1997.

Friedman, M. J., Roth, E. F., Nagel, R. L., & Trager, W. (1979). Plasmodium falciparum: physiological interactions with the human sickle cell. *Exp Parasitol, 47*(1), 73-80.

Gao, X., Starmer, J., & Martin, E. R. (2008). A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol, 32*(4), 361-369. doi: 10.1002/gepi.20310

Gething, P. W., Patil, A. P., Smith, D. L., Guerra, C. A., Elyazar, I. R., Johnston, G. L., . . . Hay, S. I. (2010). A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malar J, 10*, 378.

Gilles, H. M., Fletcher, K. A., Hendrickse, R. G., Lindner, R., Reddy, S., & Allan, N. (1967). Glucose-6-phosphate-dehydrogenase deficiency, sickling, and malaria in African children in South Western Nigeria. *Lancet, 1*(7482), 138-140.

Gilson, P. R., & Crabb, B. S. (2009). Morphology and kinetics of the three distinct phases of red blood cell invasion by Plasmodium falciparum merozoites. *Int J Parasitol, 39*(1), 91-96.

Gimnig, J. E., Kolczak, M. S., Hightower, A. W., Vulule, J. M., Schoute, E., Kamau, L., . . . Hawley, W. A. (2003). Effect of permethrin-treated bed nets on the spatial distribution of malaria vectors in western Kenya. *Am J Trop Med Hyg, 68*(4 Suppl), 115-120.

Githeko, A. K., Brandling-Bennett, A. D., Beier, M., Atieli, F., Owaga, M., & Collins, F. H. (1992). The reservoir of Plasmodium falciparum malaria in a holoendemic area of western Kenya. *Trans R Soc Trop Med Hyg, 86*(4), 355-358.

Gong, L., Maiteki-Sebuguzi, C., Rosenthal, P. J., Hubbard, A. E., Drakeley, C. J., Dorsey, G., & Greenhouse, B. (2012). Evidence for both innate and acquired mechanisms of protection from Plasmodium falciparum in children with sickle cell trait. *Blood, 119*(16), 3808-3814.

Grassi, B. (1900). *Studi di uno Zoologo Sulla Malaria.* Rome.

Grassi, B., Bignami, A., & Bastianelli, G. (1899). Ulteriore ricerche sul ciclo dei parassiti malarici umani sul corpo del zanzarone. *Atti Reale Accad Lincei, 8*, 21-28.

Grau, G. E., Taylor, T. E., Molyneux, M. E., Wirima, J. J., Vassalli, P., Hommel, M., & Lambert, P. H. (1989). Tumor necrosis factor and disease severity in children with falciparum malaria. *N Engl J Med, 320*(24), 1586-1591.

Greenwood, B., Marsh, K., & Snow, R. (1991). Why do some African children develop severe malaria? *Parasitol Today, 7*(10), 277-281. doi: 0169-4758(91)90096-7 [pii]

Guerra, C. A., Howes, R. E., Patil, A. P., Gething, P. W., Van Boeckel, T. P., Temperley, W. H., . . . Hay, S. I. (2010). The international limits and population at risk of Plasmodium vivax transmission in 2009. *PLoS Negl Trop Dis, 4*(8), e774.

Guindo, A., Fairhurst, R. M., Doumbo, O. K., Wellems, T. E., & Diallo, D. A. (2007). X-linked G6PD deficiency protects hemizygous males but not heterozygous females against severe malaria. *PLoS Med, 4*(3), e66. doi: 10.1371/journal.pmed.0040066

Gyan, B. A., Goka, B., Cvetkovic, J. T., Kurtzhals, J. L., Adabayeri, V., Perlmann, H., . . . Troye-Blomberg, M. (2004). Allelic polymorphisms in the repeat and promoter regions of the interleukin-4 gene and malaria severity in Ghanaian children. *Clin Exp Immunol, 138*(1), 145-150. doi: 10.1111/j.1365-2249.2004.02590.x

Hamblin, M. T., Thompson, E. E., & Di Rienzo, A. (2002). Complex signatures of natural selection at the Duffy blood group locus. *Am J Hum Genet, 70*(2), 369-383.

Hamel, M. J., Adazu, K., Obor, D., Sewe, M., Vulule, J., Williamson, J. M., . . . Laserson, K. F. (2011). A reversal in reductions of child mortality in western Kenya, 2003-2009. *Am J Trop Med Hyg, 85*(4), 597-605. doi: 10.4269/ajtmh.2011.10-0678

Hay, S. I., Guerra, C. A., Tatem, A. J., Noor, A. M., & Snow, R. W. (2004). The global distribution and population at risk of malaria: past, present, and future. *Lancet Infect Dis, 4*(6), 327-336.

Hay, S. I., Sinka, M. E., Okara, R. M., Kabaria, C. W., Mbithi, P. M., Tago, C. C., . . . Godfray, H. C. (2010). Developing global maps of the dominant anopheles vectors of human malaria. *PLoS Med, 7*(2), e1000209.

Hay, S. I., Smith, D. L., & Snow, R. W. (2008). Measuring malaria endemicity from intense to interrupted transmission. *Lancet Infect Dis, 8*(6), 369-378.

Hayton, K., Gaur, D., Liu, A., Takahashi, J., Henschen, B., Singh, S., . . . Wellems, T. E. (2008). Erythrocyte binding protein PfRH5 polymorphisms determine species-specific pathways of Plasmodium falciparum invasion. *Cell Host Microbe, 4*(1), 40-51.

Hill, A. V., Allsopp, C. E., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., . . . Greenwood, B. M. (1991). Common west African HLA antigens are associated with protection from severe malaria. *Nature, 352*(6336), 595-600. doi: 10.1038/352595a0

Honda, T., Miyachi, Y., & Kabashima, K. (2011). Regulatory T cells in cutaneous immune responses. *J Dermatol Sci, 63*(2), 75-82.

Howes, R. E., Patil, A. P., Piel, F. B., Nyangiri, O. A., Kabaria, C. W., Gething, P. W., . . . Hay, S. I. (2011). The global distribution of the Duffy blood group. *Nat Commun, 2*, 266.

Howes, R. E., Piel, F. B., Patil, A. P., Nyangiri, O. A., Gething, P. W., Dewi, M., . . . Hay, S. I. (2012). G6PD deficiency prevalence and estimates of affected populations in malaria endemic countries: a geostatistical model-based map. *PLoS Med, 9*(11), e1001339.

Hunja, C. W. (2006). *Frequencies of Glucose 6 Phosphate Dehydrogenase deficiency in children enrolled in a malaria vaccine clinical trial at, at Kombewa in western Kenya.* (Master of Science (Biotechnology)), Kenyatta University.

Ishino, T., Yano, K., Chinzei, Y., & Yuda, M. (2004). Cell-passage activity is required for the malarial parasite to cross the liver sinusoidal cell layer. *PLoS Biol, 2*(1), E4.

Jallow, M., Teo, Y. Y., Small, K. S., Rockett, K. A., Deloukas, P., Clark, T. G., . . . Kwiatkowski, D. P. (2009). Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nat Genet, 41*(6), 657-665. doi: 10.1038/ng.388

Jeffery, G. M. (1966). Epidemiological significance of repeated infections with homologous and heterologous strains and species of Plasmodium. *Bull World Health Organ, 35*(6), 873-882.

Joice, R., Nilsson, S. K., Montgomery, J., Dankwa, S., Egan, E., Morahan, B., . . . Marti, M. (2014). Plasmodium falciparum transmission stages accumulate in the human bone marrow. *Sci Transl Med, 6*(244), 244re245.

Kariuki, S. M., Rockett, K., Clark, T. G., Reyburn, H., Agbenyega, T., Taylor, T. E., . . . Newton, C. R. (2013). The genetic risk of acute seizures in African children with falciparum malaria. *Epilepsia, 54*(6), 990-1001. doi: 10.1111/epi.12173

Kaushal, D. C., Carter, R., Miller, L. H., & Krishna, G. (1980). Gametocytogenesis by malaria parasites in continuous culture. *Nature, 286*(5772), 490-492.

Kennedy, J. R. (2001). Malaria: a vaccine concept based on sickle haemoglobin's augmentation of an innate autoimmune process to band 3. *Int J Parasitol, 31*(11), 1275-1277.

Kifude, C. M., & Waitumbi, J. N. (2007). Abstracts of the XXth International Symposium on Technological Innovations in Laboratory Hematology, Miami, Florida, USA, 8-11 May 2007. *Int J Lab Hematol, 29 Suppl 1*, 1-112. doi: 10.1111/j.1751-553X.2007.00921.x

Kipanga, P. N., Omondi, D., Mireji, P. O., Sawa, P., Masiga, D. K., & Villinger, J. (2014). High-resolution melting analysis reveals low Plasmodium parasitaemia infections

among microscopically negative febrile patients in western Kenya. *Malar J, 13*, 429. doi: 10.1186/1475-2875-13-429

Kirk, K. (2001). Membrane transport in the malaria-infected erythrocyte. *Physiol Rev, 81*(2), 495-537.

Krishnegowda, G., Hajjar, A. M., Zhu, J., Douglass, E. J., Uematsu, S., Akira, S., . . . Gowda, D. C. (2005). Induction of proinflammatory responses in macrophages by the glycosylphosphatidylinositols of Plasmodium falciparum: cell signaling receptors, glycosylphosphatidylinositol (GPI) structural requirement, and regulation of GPI activity. *J Biol Chem, 280*(9), 8606-8616.

Kurtzhals, J. A., Adabayeri, V., Goka, B. Q., Akanmori, B. D., Oliver-Commey, J. O., Nkrumah, F. K., . . . Hviid, L. (1998). Low plasma concentrations of interleukin 10 in severe malarial anaemia compared with cerebral and uncomplicated malaria. *Lancet, 351*(9118), 1768-1772. doi: 10.1016/s0140-6736(97)09439-7

Kwiatkowski, D., Hill, A. V., Sambou, I., Twumasi, P., Castracane, J., Manogue, K. R., . . . Greenwood, B. M. (1990). TNF concentration in fatal cerebral, non-fatal cerebral, and uncomplicated Plasmodium falciparum malaria. *Lancet, 336*(8725), 1201-1204.

Laboratories, H. Haemoglobin Electrophoresis Procedure-Helena. Retrieved 10 Feb, 2015

Langhorne, J., Ndungu, F. M., Sponaas, A. M., & Marsh, K. (2008). Immunity to malaria: more questions than answers. *Nat Immunol, 9*(7), 725-732.

Lauer, S. A., Rathod, P. K., Ghori, N., & Haldar, K. (1997). A membrane network for nutrient import in red cells infected with the malaria parasite. *Science, 276*(5315), 1122-1125.

Laveran, A. (1881). Un nouveau parasite trouvé dans le sang de malades attaints de fièvre palustre. Origine parasitaire des accidents de l'impaludisme. *Bull Mém Soc Méd Hôpitaux Paris, 17*, 158-164.

Laveran, A. (1884). Traité des Fiévres Palustres avec la Description des Microbes du Paludisme. Paris.

Le Hesran, J. Y., Personne, I., Personne, P., Fievet, N., Dubois, B., Beyeme, M., . . . Deloron, P. (1999). Longitudinal study of Plasmodium falciparum infection and immune responses in infants with or without the sickle cell trait. *Int J Epidemiol, 28*(4), 793-798.

Leoratti, F. M., Farias, L., Alves, F. P., Suarez-Mutis, M. C., Coura, J. R., Kalil, J., . . . Ramasawmy, R. (2008). Variants in the toll-like receptor signaling pathway and clinical outcomes of malaria. *J Infect Dis, 198*(5), 772-780. doi: 10.1086/590440

Lettre, G., Lange, C., & Hirschhorn, J. N. (2007). Genetic model testing and statistical power in population-based association studies of quantitative traits. *Genet Epidemiol, 31*(4), 358-362. doi: 10.1002/gepi.20217

Li, C., Corraliza, I., & Langhorne, J. (1999). A defect in interleukin-10 leads to enhanced malarial disease in Plasmodium chabaudi chabaudi infection in mice. *Infect Immun, 67*(9), 4435-4442.

Liu, W., Li, Y., Shaw, K. S., Learn, G. H., Plenderleith, L. J., Malenke, J. A., . . . Sharp, P. M. (2014). African origin of the malaria parasite Plasmodium vivax. *Nat Commun, 5*, 3346.

Lukic, Z., Hausmann, S., Sebastian, S., Rucci, J., Sastri, J., Robia, S. L., . . . Campbell, E. M. (2011). TRIM5alpha associates with proteasomal subunits in cells while in complex with HIV-1 virions. *Retrovirology, 8*, 93. doi: 10.1186/1742-4690-8-93

Luzzi, G. A., Merry, A. H., Newbold, C. I., Marsh, K., Pasvol, G., & Weatherall, D. J. (1991). Surface antigen expression on Plasmodium falciparum-infected erythrocytes is modified in alpha- and beta-thalassemia. *J Exp Med, 173*(4), 785-791.

Lyke, K. E., Burges, R., Cissoko, Y., Sangare, L., Dao, M., Diarra, I., . . . Sztein, M. B. (2004). Serum levels of the proinflammatory cytokines interleukin-1 beta (IL-1beta), IL-6, IL-8, IL-10, tumor necrosis factor alpha, and IL-12(p70) in Malian children with severe Plasmodium falciparum malaria and matched uncomplicated malaria or healthy controls. *Infect Immun, 72*(10), 5630-5637.

Mackinnon, M. J., Mwangi, T. W., Snow, R. W., Marsh, K., & Williams, T. N. (2005). Heritability of malaria in Africa. *PLoS Med, 2*(12), e340. doi: 10.1371/journal.pmed.0020340

MALGEN. (2008). A global network for investigating the genomic epidemiology of malaria. *Nature, 456*(7223), 732-737.

MALGEN. (2014). Reappraisal of known malaria resistance loci in a large multicenter study. *Nat Genet, 46*(11), 1197-1204.

Manjurano, A., Clark, T. G., Nadjm, B., Mtove, G., Wangai, H., Sepulveda, N., . . . Drakeley, C. (2012). Candidate human genetic polymorphisms and severe malaria in a Tanzanian population. *PLoS One, 7*(10), e47463.

Manson, P. (1879). On the development of Filaria sanguinis hominis, and on the mosquito considered as a nurse. . *J Linn Soc Lond Zoo, 14*, 304-311.

Marquet, S., Doumbo, O., Cabantous, S., Poudiougou, B., Argiro, L., Safeukui, I., . . . Dessein, A. J. (2008). A functional promoter variant in IL12B predisposes to cerebral malaria. *Hum Mol Genet, 17*(14), 2190-2195. doi: 10.1093/hmg/ddn118

Marsh, K., English, M., Crawley, J., & Peshu, N. (1996). The pathogenesis of severe malaria in African children. *Ann Trop Med Parasitol, 90*(4), 395-402.

Marsh, K., & Kinyanjui, S. (2006). Immune effector mechanisms in malaria. *Parasite Immunol, 28*(1-2), 51-60.

Marsh, K., Otoo, L., Hayes, R. J., Carson, D. C., & Greenwood, B. M. (1989). Antibodies to blood stage antigens of Plasmodium falciparum in rural Gambians and their relation to protection against infection. *Trans R Soc Trop Med Hyg, 83*(3), 293-303.

May, J., Evans, J. A., Timmann, C., Ehmen, C., Busch, W., Thye, T., . . . Horstmann, R. D. (2007). Hemoglobin variants and disease manifestations in severe falciparum malaria. *JAMA, 297*(20), 2220-2226. doi: 10.1001/jama.297.20.2220

McCann, R. S., Ochomo, E., Bayoh, M. N., Vulule, J. M., Hamel, M. J., Gimnig, J. E., . . . Walker, E. D. (2014). Reemergence of Anopheles funestus as a vector of

Plasmodium falciparum in western Kenya after long-term implementation of insecticide-treated bed nets. *Am J Trop Med Hyg, 90*(4), 597-604.

Miller, L. H., Mason, S. J., Clyde, D. F., & McGinniss, M. H. (1976). The resistance factor to Plasmodium vivax in blacks. The Duffy-blood-group genotype, FyFy. *N Engl J Med, 295*(6), 302-304. doi: 10.1056/nejm197608052950602

Molineaux L, G. G. (1980). The Garki Project: Research on the epidemiology and control of malaria in the Sudan Savanna of West Africa. Geneva: WHO.

Mouele, R., Pambou, O., Feingold, J., & Galacteros, F. (2000). alpha-thalassemia in Bantu population from Congo-Brazzaville: its interaction with sickle cell anemia. *Hum Hered, 50*(2), 118-125. doi: 22899

Murray, C. J., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., . . . Lopez, A. D. (2012). Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet, 379*(9814), 413-431.

Nakayama, E. E., & Shioda, T. (2012). Role of Human TRIM5alpha in Intrinsic Immunity. *Front Microbiol, 3*, 97. doi: 10.3389/fmicb.2012.00097

Ndungu, F. M., Olotu, A., Mwacharo, J., Nyonda, M., Apfeld, J., Mramba, L. K., . . . Marsh, K. (2012). Memory B cells are a more reliable archive for historical antimalarial responses than plasma antibodies in no-longer exposed children. *Proc Natl Acad Sci U S A, 109*(21), 8247-8252.

Noor, A. M., Gething, P. W., Alegana, V. A., Patil, A. P., Hay, S. I., Muchiri, E., . . . Snow, R. W. (2009). The risks of malaria infection in Kenya in 2009. *BMC Infect Dis, 9*, 180.

Nussenzweig, R. S., Vanderberg, J., Most, H., & Orton, C. (1967). Protective immunity produced by the injection of x-irradiated sporozoites of plasmodium berghei. *Nature, 216*(5111), 160-162.

O'Meara, W. P., Bejon, P., Mwangi, T. W., Okiro, E. A., Peshu, N., Snow, R. W., . . . Marsh, K. (2008). Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. *Lancet, 372*(9649), 1555-1562.

Ogot, B. (1967). *History of the Southern Luo* (Vol. 1).

Ohashi, J., Naka, I., Patarapotikul, J., Hananantachai, H., Looareesuwan, S., & Tokunaga, K. (2003). A single-nucleotide substitution from C to T at position -1055 in the IL-13 promoter is associated with protection from severe malaria in Thailand. *Genes Immun, 4*(7), 528-531. doi: 10.1038/sj.gene.6364010

Omar, A. H., Yasunami, M., Yamazaki, A., Shibata, H., Ofori, M. F., Akanmori, B. D., . . . Hirayama, K. (2012). Toll-like receptor 9 (TLR9) polymorphism associated with symptomatic malaria: a cohort study. *Malar J, 11*, 168. doi: 10.1186/1475-2875-11-168

Omer, F. M., & Riley, E. M. (1998). Transforming growth factor beta production is inversely correlated with severity of murine malaria infection. *J Exp Med, 188*(1), 39-48.

Ong'echa, J. M., Raballah, E. O., Kempaiah, P. M., Anyona, S. B., Were, T., Davenport, G. C., . . . Perkins, D. J. (2011). Polymorphic variability in the 3' untranslated region (UTR) of IL12B is associated with susceptibility to severe anaemia in Kenyan children with acute Plasmodium falciparum malaria. *BMC Genet, 12*, 69. doi: 10.1186/1471-2156-12-69

Organization, W. H. (2014). Severe malaria. *Trop Med Int Health, 19 Suppl 1*, 7-131.

Parroche, P., Lauw, F. N., Goutagny, N., Latz, E., Monks, B. G., Visintin, A., . . . Golenbock, D. T. (2007). Malaria hemozoin is immunologically inert but radically enhances innate responses by presenting malaria DNA to Toll-like receptor 9. *Proc Natl Acad Sci U S A, 104*(6), 1919-1924.

Pasvol, G., Weatherall, D. J., & Wilson, R. J. (1978). Cellular mechanism for the protective effect of haemoglobin S against P. falciparum malaria. *Nature, 274*(5672), 701-703.

Perkmann, T., Winkler, H., Graninger, W., Kremsner, P. G., & Winkler, S. (2005). Circulating levels of the interleukin (IL)-4 receptor and of IL-18 in patients with Plasmodium falciparum malaria. *Cytokine, 29*(4), 153-158. doi: 10.1016/j.cyto.2004.10.008

Phawong, C., Ouma, C., Tangteerawatana, P., Thongshoob, J., Were, T., Mahakunkijcharoen, Y., . . . Khusmith, S. (2010). Haplotypes of IL12B promoter polymorphisms condition susceptibility to severe malaria and functional changes in cytokine levels in Thai adults. *Immunogenetics, 62*(6), 345-356. doi: 10.1007/s00251-010-0439-y

Piel, F. B., Howes, R. E., Patil, A. P., Nyangiri, O. A., Gething, P. W., Bhatt, S., . . . Hay, S. I. (2013). The distribution of haemoglobin C and its prevalence in newborns in Africa. *Sci Rep, 3*, 1671.

Piel, F. B., Patil, A. P., Howes, R. E., Nyangiri, O. A., Gething, P. W., Williams, T. N., . . . Hay, S. I. (2010). Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nat Commun, 1*, 104. doi: 10.1038/ncomms1104

Portugal, S., Pierce, S. K., & Crompton, P. D. (2013). Young lives lost as B cells falter: what we are learning about antibody responses in malaria. *J Immunol, 190*(7), 3039-3046.

Randall, L. M., Kenangalem, E., Lampah, D. A., Tjitra, E., Mwaikambo, E. D., Handojo, T., . . . Engwerda, C. R. (2010). A study of the TNF/LTA/LTB locus and susceptibility to severe malaria in highland papuan children and adults. *Malar J, 9*, 302. doi: 10.1186/1475-2875-9-302

Ross, P., Hall, L., Smirnov, I., & Haff, L. (1998). High level multiplex genotyping by MALDI-TOF mass spectrometry. *Nat Biotechnol, 16*(13), 1347-1351.

Ross, R. (1897). On some Peculiar Pigmented Cells Found in Two Mosquitos Fed on Malarial Blood. *Br Med J, 2*(1929), 1786-1788.

Roth, E. F., Jr., Friedman, M., Ueda, Y., Tellez, I., Trager, W., & Nagel, R. L. (1978). Sickling rates of human AS red cells infected in vitro with Plasmodium falciparum malaria. *Science, 202*(4368), 650-652.

Rowe, J. A., Moulds, J. M., Newbold, C. I., & Miller, L. H. (1997). P. falciparum rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature, 388*(6639), 292-295. doi: 10.1038/40888 [doi]

Rowe, J. A., Opi, D. H., & Williams, T. N. (2009). Blood groups and malaria: fresh insights into pathogenesis and identification of targets for intervention. *Curr Opin Hematol, 16*(6), 480-487.

Ruwende, C., Khoo, S. C., Snow, R. W., Yates, S. N., Kwiatkowski, D., Gupta, S., . . . et al. (1995). Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. *Nature, 376*(6537), 246-249. doi: 10.1038/376246a0 [doi]

Schellenberg, D., Menendez, C., Aponte, J., Guinovart, C., Mshinda, H., Tanner, M., & Alonso, P. (2004). The changing epidemiology of malaria in Ifakara Town, southern Tanzania. *Trop Med Int Health, 9*(1), 68-76.

Schmidt, N. W., Butler, N. S., Badovinac, V. P., & Harty, J. T. (2010). Extreme CD8 T cell requirements for anti-malarial liver-stage immunity following immunization with radiation attenuated sporozoites. *PLoS Pathog, 6*(7), e1000998.

Schofield, L., & Mueller, I. (2006). Clinical immunity to malaria. *Curr Mol Med, 6*(2), 205-221.

Schwartz, L., Brown, G. V., Genton, B., & Moorthy, V. S. (2012). A review of malaria vaccine clinical projects based on the WHO rainbow table. *Malar J, 11*, 11.

Shah, S. S., Macharia, A., Makale, J., Uyoga, S., Kivinen, K., Craik, R., . . . Williams, T. N. (2014). Genetic determinants of glucose-6-phosphate dehydrogenase activity in Kenya. *BMC Med Genet, 15*, 93. doi: 10.1186/s12881-014-0093-6

Sinden, R. E., & Hartley, R. H. (1985). Identification of the meiotic division of malarial parasites. *J Protozool, 32*(4), 742-744.

Singh, B., Kim Sung, L., Matusop, A., Radhakrishnan, A., Shamsul, S. S., Cox-Singh, J., . . . Conway, D. J. (2004). A large focus of naturally acquired Plasmodium knowlesi infections in human beings. *Lancet, 363*(9414), 1017-1024.

Sinka, M. E., Bangs, M. J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., . . . Hay, S. I. (2012). A global map of dominant malaria vectors. *Parasit Vectors, 5*, 69.

Smith, T., Felger, I., Tanner, M., & Beck, H. P. (1999). Premunition in Plasmodium falciparum infection: insights from the epidemiology of multiple infections. *Trans R Soc Trop Med Hyg, 93 Suppl 1*, 59-64.

Smith, T., Schellenberg, J. A., & Hayes, R. (1994). Attributable fraction estimates and case definitions for malaria in endemic areas. *Stat Med, 13*(22), 2345-2358.

Sriram, U., Xu, J., Chain, R. W., Varghese, L., Chakhtoura, M., Bennett, H. L., . . . Gallucci, S. (2014). IL-4 suppresses the responses to TLR7 and TLR9 stimulation and increases the permissiveness to retroviral infection of murine conventional dendritic cells. *PLoS One, 9*(1), e87668. doi: 10.1371/journal.pone.0087668

Stoute, J. A. (2005). Complement-regulatory proteins in severe malaria: too little or too much of a good thing? *Trends Parasitol, 21*(5), 218-223. doi: 10.1016/j.pt.2005.03.004

Sylla, E. H., Kun, J. F., & Kremsner, P. G. (2000). Mosquito distribution and entomological inoculation rates in three malaria-endemic areas in Gabon. *Trans R Soc Trop Med Hyg, 94*(6), 652-656.

Taylor, S. M., Parobek, C. M., & Fairhurst, R. M. (2012). Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis. *Lancet Infect Dis, 12*(6), 457-468.

Thathy, V., Moulds, J. M., Guyah, B., Otieno, W., & Stoute, J. A. (2005). Complement receptor 1 polymorphisms associated with resistance to severe malaria in Kenya. *Malar J, 4*, 54. doi: 10.1186/1475-2875-4-54

Timmann, C., Thye, T., Vens, M., Evans, J., May, J., Ehmen, C., . . . Horstmann, R. D. (2012). Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature, 489*(7416), 443-446.

Tiono, A. B., Guelbeogo, M. W., Sagnon, N. F., Nebie, I., Sirima, S. B., Mukhopadhyay, A., & Hamed, K. (2013). Dynamics of malaria transmission and susceptibility to clinical malaria episodes following treatment of Plasmodium falciparum asymptomatic carriers: results of a cluster-randomized study of community-wide screening and treatment, and a parallel entomology study. *BMC Infect Dis, 13*, 535. doi: 10.1186/1471-2334-13-535

Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., . . . Williams, S. M. (2009). The genetic structure and history of Africans and African Americans. *Science, 324*(5930), 1035-1044.

Traore, B., Kone, Y., Doumbo, S., Doumtabe, D., Traore, A., Crompton, P. D., . . . Pierce, S. K. (2009). The TLR9 agonist CpG fails to enhance the acquisition of Plasmodium falciparum-specific memory B cells in semi-immune adults in Mali. *Vaccine, 27*(52), 7299-7303. doi: 10.1016/j.vaccine.2009.08.023

Turner, L., Lavstsen, T., Berger, S. S., Wang, C. W., Petersen, J. E., Avril, M., . . . Theander, T. G. (2013). Severe malaria is associated with parasite binding to endothelial protein C receptor. *Nature, 498*(7455), 502-505.

Udomsangpetch, R., Sueblinvong, T., Pattanapanyasat, K., Dharmkrong-at, A., Kittikalayawong, A., & Webster, H. K. (1993). Alteration in cytoadherence and rosetting of Plasmodium falciparum-infected thalassemic red blood cells. *Blood, 82*(12), 3752-3759.

van de Vosse, E., Ali, S., de Visser, A. W., Surjadi, C., Widjaja, S., Vollaard, A. M., & van Dissel, J. T. (2005). Susceptibility to typhoid fever is associated with a polymorphism in the cystic fibrosis transmembrane conductance regulator (CFTR). *Hum Genet, 118*(1), 138-140. doi: 10.1007/s00439-005-0005-0

Verra, F., Luoni, G., Calissano, C., Troye-Blomberg, M., Perlmann, P., Perlmann, H., . . . Modiano, D. (2004). IL4-589C/T polymorphism and IgE levels in severe malaria. *Acta Trop, 90*(2), 205-209. doi: 10.1016/j.actatropica.2003.11.014

Walley, A. J., Aucan, C., Kwiatkowski, D., & Hill, A. V. (2004). Interleukin-1 gene cluster polymorphisms and susceptibility to clinical malaria in a Gambian case-control study. *Eur J Hum Genet, 12*(2), 132-138. doi: 10.1038/sj.ejhg.5201084

Walther, M., Jeffries, D., Finney, O. C., Njie, M., Ebonyi, A., Deininger, S., . . . Riley, E. M. (2009). Distinct roles for FOXP3 and FOXP3 CD4 T cells in regulating cellular immunity to uncomplicated and severe Plasmodium falciparum malaria. *PLoS Pathog, 5*(4), e1000364.

Walther, M., Woodruff, J., Edele, F., Jeffries, D., Tongren, J. E., King, E., . . . Riley, E. M. (2006). Innate immune responses to human malaria: heterogeneous cytokine responses to blood-stage Plasmodium falciparum correlate with parasitological and clinical outcomes. *J Immunol, 177*(8), 5736-5745.

Wambua, S., Mwangi, T. W., Kortok, M., Uyoga, S. M., Macharia, A. W., Mwacharo, J. K., . . . Williams, T. N. (2006). The effect of alpha+-thalassaemia on the incidence of malaria and other diseases in children living on the coast of Kenya. *PLoS Med, 3*(5), e158. doi: 10.1371/journal.pmed.0030158

Waterfall CM, C. B. (2001). Single tube genotyping of sickle cell anaemia using PCR-based SNP analysis. *Nucleic Acids Research, 29*(23), e119.

Webster, D. P., Dunachie, S., Vuola, J. M., Berthoud, T., Keating, S., Laidlaw, S. M., . . . Hill, A. V. (2005). Enhanced T cell-mediated protection against malaria in human challenges by using the recombinant poxviruses FP9 and modified vaccinia virus Ankara. *Proc Natl Acad Sci U S A, 102*(13), 4836-4841.

Weiss, G. E., Ndungu, F. M., McKittrick, N., Li, S., Kimani, D., Crompton, P. D., . . . Pierce, S. K. (2012). High efficiency human memory B cell assay and its application to studying Plasmodium falciparum-specific memory B cells in natural infections. *J Immunol Methods, 375*(1-2), 68-74.

White, N. J. (2008). The role of anti-malarial drugs in eliminating malaria. *Malar J, 7 Suppl 1*, S8. doi: 10.1186/1475-2875-7-s1-s8

WHO. (1951). Report on the malaria conference in equatorial Africa. Held under the joint auspices of the World Health Organization and of the commission for technical co-operation in Africa south of the Sahara. Kampala, Uganda, 27 November–9 December, 1950 *World Health Organ Tech Rep Ser 1951; 38: 72.* Geneva:: World Health Organization.

WHO. (2000). Severe falciparum malaria. *Trans R Soc Trop Med Hyg, 94*(supplement 1).

WHO. (2010). Guidelines for the treatment of malaria (Second edition ed.).

WHO. (2013). World Malaria Report Geneva, Switzerland: World Health Organization.

Williams, T. N. (2006). Red blood cell defects and malaria. *Mol Biochem Parasitol, 149*(2), 121-127. doi: 10.1016/j.molbiopara.2006.05.007

Williams, T. N., Maitland, K., Bennett, S., Ganczakowski, M., Peto, T. E., Newbold, C. I., . . . Clegg, J. B. (1996). High incidence of malaria in alpha-thalassaemic children. *Nature, 383*(6600), 522-525. doi: 10.1038/383522a0

Williams, T. N., Mwangi, T. W., Roberts, D. J., Alexander, N. D., Weatherall, D. J., Wambua, S., . . . Marsh, K. (2005). An immune basis for malaria protection by the sickle cell trait. *PLoS Med, 2*(5), e128. doi: 10.1371/journal.pmed.0020128

Williams, T. N., Mwangi, T. W., Wambua, S., Alexander, N. D., Kortok, M., Snow, R. W., & Marsh, K. (2005). Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases. *J Infect Dis, 192*(1), 178-186. doi: JID34144 [pii] 10.1086/430744 [doi]

Williams, T. N., Mwangi, T. W., Wambua, S., Peto, T. E., Weatherall, D. J., Gupta, S., . . . Marsh, K. (2005). Negative epistasis between the malaria-protective effects of alpha+-thalassemia and the sickle cell trait. *Nat Genet, 37*(11), 1253-1257. doi: 10.1038/ng1660

Wilson, J. N., Rockett, K., Jallow, M., Pinder, M., Sisay-Joof, F., Newport, M., . . . Kwiatkowski, D. (2005). Analysis of IL10 haplotypic associations with severe malaria. *Genes Immun, 6*(6), 462-466.

Wipasa, J., Suphavilai, C., Okell, L. C., Cook, J., Corran, P. H., Thaikla, K., . . . Hafalla, J. C. (2010). Long-lived antibody and B Cell memory responses to the human malaria parasites, Plasmodium falciparum and Plasmodium vivax. *PLoS Pathog, 6*(2), e1000770.

Woolhouse, M. E. (1998). Patterns in parasite epidemiology: the peak shift. *Parasitol Today, 14*(10), 428-434.

World Health Organization, C. D. C. (2000). Severe falciparum malaria. . *Trans R Soc Trop Med Hyg, 94 Suppl 1*, S1-90.

Yuthavong, Y., Butthep, P., Bunyaratvej, A., Fucharoen, S., & Khusmith, S. (1988). Impaired parasite growth and increased susceptibility to phagocytosis of Plasmodium falciparum infected alpha-thalassemia or hemoglobin Constant Spring red blood cells. *Am J Clin Pathol, 89*(4), 521-525.

Zenewicz, L. A., & Flavell, R. A. (2011). Recent advances in IL-22 biology. *Int Immunol, 23*(3), 159-163. doi: 10.1093/intimm/dxr001

Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W., & Arnheim, N. (1992). Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A, 89*(13), 5847-5851.

Zurawski, G., & de Vries, J. E. (1994). Interleukin 13, an interleukin 4-like cytokine that acts on monocytes and B cells, but not on T cells. *Immunol Today, 15*(1), 19-26. doi: 10.1016/0167-5699(94)90021-3

APPENDIX I: Informed consent form for parent/guardian in the study

FK grade level: 8.7
Title: The effect of inherited blood conditions on the risk of malaria.

Consent version number 4.
Original dated 27 September 2011

Investigators:
KEMRI/CDC: Godfrey Bigogo (Principal Investigator), Oscar Nyangiri, Deron Burton, Robert Breiman, Simon Kariuki, Meghna Desai, Joel Montgomery, Frank Odhiambo, Kayla Laserson, John Vulule
KEMRI/Wellcome Trust: Thomas Williams, Alexander Macharia, Carolyne Ndila
CDC-Atlanta: Craig Hooper

Purpose of research:
Children inherit some red blood cell conditions, which affect the progress of malaria. One example of such conditions is sickle cell anaemia. In this disease, the red blood cells have abnormal shapes and do not work properly. This causes anaemia. We would like to find out how these blood conditions affect malaria in the children of Asembo. To do this, we would like to test for the inherited blood conditions and follow up the children to find out how they affect their health. We hope this will help us understand malaria so that we can give suggestions for prevention and treatment.

What the research involves:
We are inviting all children less than six months and their siblings up to the age of 12 years, who are born to mothers in the IEIP health study to participate in this study. We will enrol infants and their siblings within two years from the start of the study. We would like to ask your child to be part of this study. This study will involve two steps: testing for inherited blood conditions, and follow up to check for their health effects. To test blood, we will need a sample as soon as possible after birth. This will allow us to see the health effects of the blood conditions early.

We will make a small prick, like the one normally used to test for malaria on the heel of the child. We will collect less than half a teaspoon of blood from this prick to test for the inherited blood conditions. We may collect the sample in one of the following ways:

1) For those born in the hospital, the sample may be collected in the hospital at birth.

2) For those born at home: The sample may be collected through the EMEP pregnancy outcome follow up visit at home within a few days of birth. This visit checks the health of the child for the women participating in EMEP.

3) A nurse can also visit the home or you can bring the child to Lwak Hospital within 6 months of birth to give the sample.

4) Some mothers are already participating in the study that monitors effectiveness of fansidar in pregnancy, (the IPTp-Mon) study. This study collects blood from the cord of the baby when they are born to test if they have malaria. If you agree, we will test this blood for the inherited blood conditions instead of collecting blood from the baby. The advantage of this is that the child will not feel any pain for blood collection. If you are not participating in the IPTp-Mon study, we may still collect cord blood to test for these inherited blood conditions if you agree.

We will test for sickle and other inherited blood conditions at KEMRI and the Wellcome Trust Centre for Human Genetics in Oxford, United Kingdom. We may do tests to confirm our results at the Centers for Disease Control in the United States. If we find children with sickle cell disease, we will provide the test result and health education to their parents.

Follow up: Your child is part of the International Emerging Infections Program (IEIP) health follow up. This is a home and clinic health follows up. At home, this follow up collects sickness information every two weeks to look for common diseases such as malaria. At Lwak health centre, the study records all visits. We will get information about the health of your child from this follow up. Follow up will go on until 2015. It may go on for a longer time. We will look at this information to find the relation between the blood condition and your child's health.

Some children have malaria parasites but do not show symptoms of illness. To determine which level of parasites cause malaria illness, your child will be requested to give blood to test for malaria during the forthcoming June/July malaria transmission season in Lwak hospital. Siblings of the child up to the age of 12 years old will also be invited to give a

sample to test for red cell genetics and malaria during the June/July malaria transmission season.

Benefits and risks:

There are some benefits and few risks in this study. The test results will show whether your child has sickle cell disease or not. This will allow you to take care of the child properly. Children who are found to have sickle cell Anaemia in this study will be offered outpatient tests and medicines that prevent some sickle illnesses every 3 months until they are two years old. These outpatient tests and medicines will be at Lwak hospital and will consist of standard diagnostic, therapeutic and preventive measures such as a full blood count test, folic acid, penicillin and palludrin. Their siblings who have the sickle symptoms will also be offered diagnosis and treatment for the same period. This study will also pay for health insurance with National Hospital Insurance Fund (NHIF) for the family until the affected child attains 2 years of age. In addition, during the cross sectional survey of June/July, the child will be tested for malaria and if found positive treatment will be given.

We will not pay you for being in this study. However, we will refund 200 Kenyan shillings as transport for those who bring their children to Lwak health centre to give a blood sample for this study within 6 months. During the cross sectional survey, parents who bring their children to Lwak hospital will also be reimbursed.

There could be a little pain and bleeding from the heel prick. This is the usual pain felt when you give a blood sample in hospital. In rare cases, swelling and infection may occur, which we will aim to prevent. If your child has any of these symptoms, they will most likely go away by themselves in two weeks.

Privacy and Confidentiality:

There is a rare chance that others will find out about the child's personal information. The chances of this are low because the child's records for this study will have an identification number, but not the name. Any information or samples that leave the hospital will have only this number. Your child's name will not be on it. In case of sickle cell Anaemia, the child will need treatment. In such a case, only the doctors and nurses treating your child and the person in charge of the study at the hospital will be able to see your child's name. We will only share the results of the study, without names, with others. We will keep the information

in password-protected computers. Only the study staff will have the password. This will reduce the chance of other people finding out the information of your child.

Your Right to Participate, Say No, or Withdraw:

It is up to you to decide if you want to take part. You have the right to refuse. If you decide to take part, we will ask you to sign two copies of the consent form. One copy will remain in the study file and we will give you the other copy to keep. If you do not want to continue with the study, you can stop at any time. To do this, please inform Godfrey Bigogo at KEMRI/CDC (phone number: +254-57-2022902/59/83).

Injury as a result of research:

This will be a routine blood draw, so no special injuries are expected. If your child has an injury or a bad side effect as a result of being in this study, s/he will receive medical care or treatment for the injury, but you will not receive payment or compensation. If you believe your child has been injured or have questions about a research-related injury, please contact Godfrey Bigogo at KEMRI/CDC (phone number: +254-57-2022902/59/83)."

If you have any questions about this study, please contact:

Godfrey Bigogo at KEMRI/CDC (phone number: +254-57-2022902/59/83).

If you have any questions about your child's rights in this study, please contact:

The Secretary or Chairman of the KEMRI Ethical Review Committee,

PO Box 54840 00100, Nairobi, Kenya:

Telephone: +254-20-2722541

Email: ERC@kemri.org.

These phone numbers are not for medical emergencies. If you or your child has a medical emergency, please go to the nearest health facility.

Consent for study participation:

Participation in this study does not mean participation in all studies for mothers and their babies. You are free to participate in one activity without participating in another.

I have read or had this consent form read to me. I have had a chance to ask questions. I understand that I am free to choose not to be in this study. Saying "No" will have no effect on me or my family member's health care. With my signature or thumbprint below, I give my voluntary consent to have my child or family member participate in this research study.

_____ _____

Name of Child **Child's Date of Birth (DD/MM/YY)**

_____ _____

Name of Parent/Responsible Adult **Relationship to child**

_____ _____

Signature/Thumbprint of Parent/Responsible Adult **Date (DD/MM/YY)**

_____ _____

Signature of Witness (for illiterate parents/guardians) **Date (DD/MM/YY)**

Permission for Storage, Shipping and future use of samples:

We would like to store the samples in this study. This will allow us to do other infectious disease genetic tests in future. The samples will be stored indefinitely at KEMRI/CDC in Kisumu and KEMRI/Wellcome Trust in Kilifi. If you agree, we will ship samples to KEMRI-Wellcome Trust in Kilifi, the Wellcome Trust Centre for Human Genetics in Oxford, UK or the CDC in Atlanta for tests.

Guardian/Parent

Name of child	**Signature/thumbprint**	**Date (DD/MM/YY)**
_____	_____	_____

Permission to link information from other KEMRI/CDC surveillance projects:

KEMRI/CDC gathers population and health information regularly in areas near Lwak Hospital. You or your family member may already take part in one of these ongoing surveillance projects. We will link data about you or your family member that we collect as part of this study to data from other KEMRI/CDC projects in your area, if you agree.

Guardian/Parent

Name of child	**Signature/thumbprint**	**Date (DD/MM/YY)**
_____	_____	_____

I have explained this form to the person above. I have answered any questions about this study.

_____

Name of Investigator/Study Staff **Date (DD/MM/YY)**

_____

Signature of Investigator/Study Staff **Date (DD/MM/YY)**

APPENDIX II: Reagents and solutions

1. Cross sectional survey

10% Giemsa stain

EDTA- Ethylene Diamine Tetra Acetic Acid

2. Haemoglobin phenotyping

Cellulose Acetate electrophoresis- Helena systems Kit

3. ABO phenotyping

Antisera for blood group A, B and Rhesus Blood group

4. DNA extraction

QIA amp Blood Minikit – Qiagen

Protein K

Phosphate buffered Saline – PBS

Buffer AL

Ethanol 96-100%

Wash Buffer 1 (AW-1)

Elution Buffer

Wash Buffer 2 (AW-2)

Elution Buffer (Buffer AE or Distilled water)

5. Typing for the α^+ thalassaemia deletion

Qiagen PCR Kit

PCR Master Mix

RNase free water

5X Q solution

2X fast cycling PCR master mix

10x coral load cycling dye

Alpha Thalassemia Specific primers

Primer 376

Primer 377

Primer 378

LIS1-F

LIS1-R

Controls

Homozygous Alpha Thalassemia Control

Normal Control

No Template Control

1% Agarose gel

6. Multiplex genotyping

Pico Green® Reagent – (Invitrogen, Paisley, UK)

PCR Master Final mix Volume

1:10 Diluted N15 Primer

8Mm Pooled d NTPS (Sigma – Aldrich, UK)

50Mm Magnesium Chloride – Mg Cl2 (Bioline, UK)

10 x Bio Taq Buffer (Bioline, UK)

5U/µL Biotaq Polymerase (Bioline, UK)

MilliQ Water

Genomic DNA (g DNA)

2 % Agarose gel

7. Sequenom mass array genotyping platform

iPLEX Design – Universal 10 base 5' Sequence

- 20 bases of sequence specific bases
- CEPH and YRI Hap MAP DNAs

Sample preparation

- Phenol red Solution 0.1mg /ml
- i PLEX primers – Metabion International AG (Martinsried , Germany)

First Round Reaction Master mix

50 mM Magnesium Chloride (Mgcl2)

dNTPs

10X Hot Star Taq Buffer (Qiagen)

Hot Star Taq (5 u/μL) Qiagen

Milli Q Water

First Round Reaction

PCR Master Mix

2% Agarose gel

8. Shrimp-alkaline phosphatase treatment

iPLEX Shrimp Alkaline phosphatase (SAP)

Primer Extension Reaction

Extension Taq

Extension Buffer

Primers 300 Mm

Ion exchange resin

Milli Q Water

BUFFERS

DNA extraction

1. Wash Buffer 1 (AW-1)

2. Wash Buffer 2 (AW-2)
3. Elution Buffer (Buffer AE, or distilled water)
4. QIA amp DNA Extraction solution Buffer- AL – Guanidine Chloride +
5. QIA amp DNA Buffer AE- 50-01-1 Guanidinium Chloride 20-50%

Typing for the α^+ thalassaemia deletion

First Round reaction mix

- Hot star Taq Buffer

EQUIPMENT

Glass slides

Care start Pf HRP2 KIT- RDT Conduct of cross sectional survey

Rods

Pipettes

Micro centrifuge tubes

Flat Cap Strips (Thermo Fischer Scientific)

Microseal 'A' lids (Bio – Rad)

Spectrochips

Microscopes

Refrigerators

Centrifuges

QIA amp DNA blood Mini kit – (Qiagen)

Ultraviolet Ray Trans illuminator- (Gel Doc)

Mass Spectrometer

Qiagen PCR kit

384 well - PCR Plate – Thermo Fischer

96 well plates (Thermo- fast ® 96 – skirted Thermo Fischer Scientific , UK)

APPENDIX III: SNPs assayed for in the multiplex Sequenom platform

Reference SNP number	Gene symbol	Strand	Reference allele	Derived allele	Single letter SNP code
rs1803632	GBP7	1	G	C	S
rs2814778	DARC	-1	A	G	R
rs3024500	IL10	1	A	G	R
rs1800896	IL10	1	T	C	Y
rs1800890	IL10	1	A	T	W
rs17047660	CR1	1	A	G	R
rs17047661	CR1	1	A	G	R
rs17561	IL1A	-1	G	T	K
rs1143634	IL1B	-1	C	T	Y
rs708567	IL17RE	-1	G	A	R
rs352140	TLR9	-1	A	G	R
rs187084	TLR9	-1	C	T	Y
rs6780995	IL17RD	1	G	A	R
rs4833095	TLR1	1	C	T	Y
rs5743611	TLR1	-1	G	C	S
rs5743810	TLR6	-1	C	T	Y
rs5743809	TLR6	-1	T	C	Y
rs1801033	C6	-1	A	C	M
rs2706384	IRF1	-1	C	A	M
rs20541	IL13	-1	C	T	Y
rs2243250	IL4	1	C	T	Y
rs2239704	LTA	-1	G	T	K
rs909253	LTA	-1	T	C	Y
rs1799964	TNF	1	T	C	Y
rs1800750	TNF	1	G	A	R
rs1800629	TNF	1	G	A	R
rs361525	TNF	1	G	A	R
rs3093662	TNF	1	A	G	R
rs2242665	CTL4	-1	G	A	R

rs1555498	IL20RA	1	C	T	Y
rs2075820	NOD1	-1	G	A	R
rs3211938	CD36	1	T	G	K
hCD36_G1439C	CD36	1	G	C	S
rs17140229	CFTR	1	T	C	Y
rs4986790	TLR4	1	A	G	R
rs4986791	TLR4	1	C	T	Y
rs8176746	ABO	-1	C	A	M
rs8176719	ABO	-1	I	D	I
rs33950507	HBB	-1	G	A	D
rs334	HBB	-1	A	T	W
rs33930165	HBB	1	G	A	R
rs7935564	TRIM5	1	G	A	R
rs542998	RTN3	1	T	C	Y
rs2227507	IL22	1	T	C	Y
rs1012356	IL22	1	A	T	W
rs2227491	IL22	1	T	C	Y
rs2227485	IL22	1	G	A	R
rs2227478	IL22	1	G	A	R
rs229587	SPTB	1	T	C	Y
rs2230739	ADCY9	-1	A	G	R
rs10775349	ADCY9	-1	C	G	S
rs1805015	IL4R	1	T	C	Y
rs2535611	ADORA2B	1	T	C	Y
rs2297518	NOS2	1	G	A	R
rs1800482	NOS2	-1	G	C	S
rs9282799	NOS2	-1	C	T	Y
rs8078340	NOS2	-1	C	T	Y
rs373533	EMR1	-1	G	T	K
rs461645	EMR1	-1	T	C	Y
rs1799969	ICAM1	1	G	A	R
rs5498	ICAM1	1	A	G	R
rs8386	GNAS	1	C	T	Y
rs1128127	DERL3	1	G	A	R
rs3092945	CD40LG	1	T	C	Y

rs1126535	CD40LG	1	T	C	Y
rs1050829	G6PD	1	T	C	Y
rs1050828	G6PD	1	C	T	Y

APPENDIX IV: Multiplex genotyping at the Wellcome Trust Centre for Human Genetics, Oxford

Multiplex genotyping

Genotyping for multiple SNPs in the sample was done using a multiplex method at the Wellcome Trust Centre for Human Genetics in Oxford, UK. DNA concentrations were measured using the PicoGreen® reagent (Invitrogen, Paisley, UK), which uses fluorophores that become fluorescent upon binding to DNA. Fluorescence intensity is directly proportional to the amount of DNA in the sample allowing quantification of the amount of DNA in the sample. Genomic DNA (gDNA) was then amplified by Primer Extension Pre-amplification (PEP) (Zhang *et al.*, 1992) before genotyping for candidate gene polymorphisms using the Sequenom Mass Array™ platform [25, 26]. PEP allows the amplification of whole genomic DNA from as little as a single cell and is an important genomic DNA enrichment step before Sequenom Mass Array™ genotyping.

Genomic DNA was then diluted to 1ng/μL in 96-well plates (Thermo-Fast® 96-skirted, Thermo Fisher Scientific, UK), leaving 2 to 3 empty wells for water controls. A PCR master mix final volume of 45μL was made consisting of the following: 2.2μL of 1:10 diluted N15 primers (Genetix Ltd, UK); http://www.genetix.com, 1.25μL 8mM pooled dNTPs (Sigma-Aldrich, UK), 2.5μL 50mM $MgCl_2$ (Bioline, UK); http://www.bioline.com, 5 μL of 10X BioTaq buffer (Bioline, UK), 0.5μL 5U/μL Biotaq polymerase (Bioline) and 33.55μl MilliQ water was added to each of the 96 wells. 5μL of gDNA (1ng/μL) was added to the PEP master mix in each well before sealing the plates with flat cap strips (Thermo Fisher Scientific) and thermo cycling. The PCR thermo cycling programme used was as follows:

94^{0}C for 3 min; 50 cycles of (94^{0}C for 1min, 37^{0}C for 2 min, Ramp to 55^{0}C at 0.1/sec, 55^{0}C for 4 min) and a final extension of 72^{0}C for 5 min, maintain at 4^{0}C.

After amplification, the PEP DNA was stored neat at -20^{0}C until it was required.

Before genotyping on the Sequenom platform, the PEP DNA was tested by selecting 12 samples at random from the PEP plate, amplifying as described below for Sequenom genotyping (with modifications) and running the samples on 2% agarose gels to check band intensity and fidelity. Modifications on the Sequenom genotyping are described below:

A final reaction volume of 20μL; 1μL of neat PEP and a single primer pair designed from a current iPLEX assay as follows:

forward primer: ACGTTGGATGTCTGTAGTGATGGAGGGATG

reverse primer: ACGTTGGATGGTGTCCTCTCCCTTGTAAAC

Sequenom Mass Array genotyping platform

We used a multiplex genotyping platform (P. Ross *et al.*, 1998; Wilson *et al.*, 2005) to type for multiple SNPs simultaneously. The specific kit used was the SEQUENOM® iPLEX® Gold which allows up to 40 SNPs to be designed into a single reaction and up to 384 samples to be processed on a single chip.

SNP sets: Genotyping was performed for SNP sets chosen from literature searches of SNPs showing associations with malaria or infectious diseases. See APPENDIX H for a list of the

SNPs and associated literature. In addition, SNPs intended for determination of gender by the amelogenin gene differences in the X and Y chromosomes were typed (Eng *et al.*, 1994).

iPLEX design: The SEQUENOM® assay design process (**www.sequenom.com**) requires polymorphism sequences. These were downloaded from Ensembl (**http://www.ensembl.org**) and formatted appropriately for the assay design. SEQUENOM® RealSNP™ Assay Database (**http://www.realsnp.com**) tools ProxSNP and PreEXTEND were then used for identification of the target SNPs and in the design of first round PCR primers. The MassARRAY® Assay Design v3.1 Software was used for Multiplex design. iPLEX methodology was employed where common settings for assay design included addition of a universal 10 base 5' sequence and at least 20 bases of sequence specific bases. The design ensured amplicon sizes ranging between 80 to 120 bases, and on average the amplicon size was 100 bases. For universal extension primers, the target mass range was 4500 to 10,000Da (~15-mer to ~29-mer oligos). 2 multiplexes were designed as above then tested using a panel of CEPH and YRI HapMap DNAs. This allowed identification of assays that were either poorly performing or that had poor concordance for removal from the multiplex system.

Sample preparation: PEP DNA samples were thawed before dilution at 1:10 using a phenol red solution (0.01mg/ml) to aid tracking into 384-well plates. Diluted PEP for use for the assay was kept at 4^0C for a maximum of 2 days and was not freeze thawed more than twice. All the unused diluted PEP was frozen at -20^0C. iPLEX primers were purchased lyophilised from Metabion International AG (Martinsried, Germany). First round primers and extension primers were hydrated to 100μM and 300μM respectively before storage at -20^0C.

First-Round reaction master mix: For each 384 well plate a master mix comprising of the following was prepared, allowing some extra volume:

3.3µL of each first-round primer (100mM), 214.5µL MgCl2 (50mM), 66µL dNTPs (25mM pooled), 412.5µL 10X HotStar Taq buffer (Qiagen), 132µL HotStar Taq (5U/ µL) (Qiagen) and milliQ water to make a final volume of 1980µL.

First-Round Reaction: 4.5µL per well of PCR master mix was plated into a 384-well PCR plate (Thermo Fisher Scientific). To each well, 3µL of 1:10 diluted PEP DNA was added. The plate was sealed using Microseal 'A' lids (Bio-Rad) and thermo cycling was done using the following conditions: 94^0C for 15 min, 44 cycles of (94^0C for 20 sec, 56^0C for 30 sec, 72^0C for 1 min) and a final extension of 72^0C for 3 min before maintenance at 4^0C. To confirm that PCR was successful a 1µL sample from each well of a single row was run on a 2% agarose gel.

Shrimp alkaline phosphatise treatment: All dNTP's that were not incorporated into the PCR reaction above were destroyed by adding 2 µL of iPLEX shrimp-alkaline phosphatise (SAP) mixture to the first-round PCR reaction mixture and incubating at 37^0C for 40 min. This was followed by 5 minute denaturation step at 85^0C and a cooling step of 15^0C for 15 min.

Primer extension reaction: Final reaction concentrations for the extension primers were dependent on their molecular mass based on SEQUENOM® protocol guidelines as follows:

< 5800Da 0.84 μM, 5800 to 7000Da 1.04μM, 7000 to 10,000Da 1.25 μM, >10,000Da 1.5μM.

To allow for primer extension in the sample plate, 2μL of a mixture containing 0.2μL iPLEX termination mixture, 0.041μL extension Taq, 0.2μL extension buffer and primers (300mM) according to the molecular masses as follows; 0.025μL per primer up to 5800Da, 0.0312 μL per primer for 5800 to 7000Da, 0.0375 μL per primer for 7000 to 10,000 Da, 0.045 μL per primer for >10,000 Da. The final extension volume was 9μL (5ul first-round reaction, 2ul SAP and 2ul of extension mixture).

Extension thermo cycling was performed on an MJ Tetrad as follows:

94^0C for 30 sec, 40 cycles of (94^0C for 5 sec, 5 cycles of (52^0C for 5 sec, 80^0C for 5 sec)) then 72^0C for 3 min and 15^0C for 15min.

After the reaction, 6mg ion-exchange resin and 16μL MilliQ water was added per well. The plates were sealed, rotated for 30 min and centrifuged to pellet the resin. Samples were then 'spotted" onto SpectroCHIPS and run on the Mass-Spectrometer.

Data was checked for consistency and genotypes using the SEQUENOM® Typer 4.03 software. Data were downloaded and stored on a central database for any further curation. Genotype data were maintained with reference to the sequence strand used for the design process.

APPENDIX V: Scientific and Ethical Review Board Approvals

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840 - 00200 NAIROBI, Kenya
Tel: (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

ESACIPAC/SSC/9521 11th July, 2011

Godfrey Bigogo

Thro'
Director, CGHR
KISUMU

FORWARDED
DIRECTOR
CENTRE FOR GLOBAL HEALTH RESEARCH

REF: SSC: No. 2074 (Revised) –Cohort study of human genetic factors and childhood malaria in the Asembo area of Western Kenya

Thank you for your letter dated 1st July, 2011 responding to the comments raised by the KEMRI SSC.

I am pleased to inform you that your protocol now has formal scientific approval from SSC.

The SSC however, advises that work on the proposed study can only start after ERC approval.

Sammy Njenga, PhD
SECRETARY, SSC

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **October 18, 2011**

TO: **GODFREY BIGOGO (PRINCIPAL INVESTIGATOR)**

THROUGH : **DR. JOHN VULULE,**
DIRECTOR, CGHR
KISUMU

FORWARDED
DIRECTOR
CENTRE FOR GLOBAL HEALTH RESEARCH

Dear Sir,

RE: SSC PROTOCOL No. 2074 – REVISED (*RE-SUBMISSION*): COHORT STUDY OF HUMAN GENETIC FACTORS AND CHILDHOOD MALARIA IN THE ASEMBO AREA OF WESTERN KENYA

We acknowledge receipt of:

a. The Revised Study Protocol – version dated 15 August 2011;
b. The Revised Informed Consent Documents – English, Kiswahili and Dholuo versions
c. The separate consent document for sample/data collection, storage and future use

This is to inform you that the Ethics Review Committee (ERC) reviewed the documents listed above and is satisfied that the issues raised at the initial review have been adequately addressed.

The study is granted approval for implementation effective this **18th day of October 2011**. Please note that authorization to conduct this study will automatically expire on **October 17, 2012**. If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to the ERC Secretariat by **August 3, 2012**.

Any unanticipated problems resulting from the implementation of this protocol should be brought to the attention of the ERC. You are also required to submit any proposed changes to this protocol to the ERC to initiation and advise the ERC when the study is completed or discontinued.
You may embark on the study.

Sincerely,

Caroline Kithinji,
FOR: SECRETARY,
KEMRI/NATIONAL ETHICS REVIEW COMMITTEE

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **April 20, 2012**

TO: **GODFREY BIGOGO (PRINCIPAL INVESTIGATOR)**

THROUGH : **DR. JOHN VULULE,**
DIRECTOR, CGHR
KISUMU

24/4/2012

Dear Sir,

RE: **SSC PROTOCOL No. 2074 – REVISED (*RE-SUBMISSION*): COHORT STUDY OF HUMAN GENETIC FACTORS AND CHILDHOOD MALARIA IN THE ASEMBO AREA OF WESTERN KENYA**

This is to inform you that at the 200th meeting of the KEMRI Ethics Review Committee held on April 17, 2012, the request for amendment for the above referenced research proposal was discussed.

The following amendments are proposed:

1. A description of the type of care to be provided to children who are diagnosed with sickle cell anemia until they attain 2 years of age.
2. Additional counseling and education on the condition and indications for seeking care and referral will be offered to the parents of the affected children.

The Committee concluded that the amendments are justified and do not alter the risk/benefit status of the study and therefore granted approval for implementation. You are also required to submit any further requests for changes to the approved protocol to the SSC and ERC prior to initiation.

Sincerely,

DR. CHRISTINE WASUNNA,
ACTING SECRETARY,
KEMRI ETHICS REVIEW COMMITTEE

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **July 3, 2012**

TO: **DR. GODFREY BIGOGO (PRINCIPAL INVESTIGATOR)**

THROUGH: **DR. JOHN VULULE,**
DIRECTOR, CGHR
KISUMU

Dear Sir,

RE: **SSC PROTOCOL No. 2074 (*RE-SUBMISSION: REQUEST FOR AMENDMENT 2*): COHORT STUDY OF HUMAN GENETIC FACTORS AND CHILDHOOD MALARIA IN THE ASEMBO AREA OF WESTERN KENYA (*VERSION 3 DATED JUNE 2012*)**

Reference is made to your letter dated 26 June 2012. The ERC Secretariat acknowledges receipt of your application on 28th June 2012.

This is to inform you that the Committee notes that more mothers enrolled in the study favour cord blood collection over heel prick collection and determines that the issue raised at the 202nd ERC meeting 28th May 2012 is adequately addressed. You are therefore authorized to implement Amendment 2 accordingly.

You are required to submit any further changes to this study to the SSC and ERC for review and the changes should not be initiated until written approval from the ERC is received. Please note that any unanticipated problems resulting from the conduct of this study should be brought to the attention of the ERC and you should advise the ERC when the study is completed or discontinued.

You may continue with the study.

Sincerely,

DR. CHRISTINE WASUNNA,
ACTING SECRETARY,
KEMRI ETHICS REVIEW COMMITTEE

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **May 30, 2013**

TO: **DR. GODFREY BIGOGO**
PRINCIPAL INVESTIGATOR

THROUGH: **DR. JOHN VULULE,**
DIRECTOR, CGHR
KISUMU

Dear Sir,

RE: SSC PROTOCOL No. 2074 (*REQUEST FOR AMENDMENT 3*): COHORT STUDY OF HUMAN GENETIC FACTORS AND CHILDHOOD MALARIA IN THE ASEMBO AREA OF WESTERN KENYA

This is to inform you that at the 215th meeting of the KEMRI Ethics Review Committee held on 21st May 2013, the request for amendment to the above referenced research proposal was discussed.

The Committee noted that:

(a) The investigators are requesting to amendment the recruitment criteria of study participants as enrolment has been slower than expected. The investigators request to enroll any siblings (up to the age of 12 years) of both current and future recruits to the study.

(b) To use logistic regression approach that will require data from asymptomatic patients in the community as well as sick children during illness episodes.

The Committee concluded that the suggested amendments are justified and consequently granted approval for implementation. You are required to submit any further requests for changes to this version of the protocol to the SSC and ERC for review and approval prior to implementing any additional changes.

Yours sincerely,

EAB

DR. ELIZABETH BUKUSI,
ACTING SECRETARY,
KEMRI ETHICS REVIEW COMMITTEE

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **September 21, 2012**

TO: **MR. GODFREY BIGOGO (PRINCIPAL INVESTIGATOR)**

THROUGH: **DR. JOHN VULULE,**
THE DIRECTOR, CGHR,
KISUMU

FORWARDED
DIRECTOR
CENTRE FOR GLOBAL HEALTH RESEARCH

Dear Sir,

RE: **SSC PROTOCOL No. 2074 (*REQUEST FOR ANNUAL RENEWAL):* HUMAN GENETIC FACTORS AND CHILDHOOD MALARIA IN THE ASEMBO AREA IN WESTERN KENYA**

The ERC Secretariat acknowledges receipt of the following documents;

1. Continuing Review Report for the period to **18th October 2011** to **30th July 2012.**
2. Study Protocol Version No.03 dated June 2012.
3. Informed Consent Documents in English, Dholuo and Kiswahili - Version No.03 dated June 2012.

This is to inform you that at the 207th meeting of the KEMRI Ethics Review Committee held on 05th September 2012, the Committee reviewed the above referenced application and made note of the following:

(a) Two hundred and twenty one (221) children have been enrolled into the study out of the total targeted of1600 children.
(b) Preliminary analysis of genotypic data is ongoing.
(c) The planned activity for the next project period is to continue with recruitment, genotyping and surveillance for disease events, data analysis and report writing.

The study was granted approval for continuation effective the **5th day of September 2012.** Please note that authorization to conduct this study will automatically expire on **September 4, 2013.** If you plan to continue with data collection or analysis beyond this date, please submit an application for continuing approval to the ERC Secretariat by **July 24, 2013.**

You are required to submit any proposed changes to this study to the SSC and ERC for review and the changes should not be initiated until written approval from the ERC is received. Please note that any unanticipated problems resulting from the conduct of this study should be brought to the attention of the ERC and you should advise the ERC when the study is completed or discontinued.

You may continue with the study.

DR. CHRISTINE WASUNNA,
ACTING SECRETARY,
KEMRI ETHICS REVIEW COMMITTEE

KENYA MEDICAL RESEARCH INSTITUTE
CENTRE FOR GLOBAL HEALTH RESEARCH
18 OCT 2013
P.O. Box 1578, KISUMU

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **October 11, 2013**

TO: **MR. GODFREY BIGOGO,**
PRINCIPAL INVESTIGATOR

THROUGH: **DR. STEPHEN MUNGA,**
ACTING DIRECTOR, CGHR
KISUMU

FORWARDED
DIRECTOR
CENTRE FOR GLOBAL HEALTH RESEARCH

Dear Sir,

RE: SSC PROTOCOL No. 2074: RESUBMISSION (*REQUEST FOR ANNUAL RENEWAL):* HUMAN GENETIC FACTORS AND CHILDHOOD MALARIA IN THE ASEMBO AREA OF WESTERN KENYA

Reference is made to your letter dated 8th October, 2013. The ERC Secretariat acknowledges receipt of the revised annual report on October 10, 2013.

This is to inform you that the Ethics Review Committee (ERC) reviewed the document listed above and approved the application for another year.

This approval is valid from today, **11th October 2013** through to **11th October 2014.** Please note that authorization to conduct this study will automatically expire on **10th October 2014.** If you plan to continue with data collection or analysis beyond this date please submit an application for continuing approval to the **ERC** secretariat by **29th August 2014.**

You are required to submit any amendments to this protocol and other information pertinent to human participation in this study to the SSC and ERC for review prior to initiation.

Yours faithfully,

EAB

DR. ELIZABETH BUKUSI,
ACTING SECRETARY,
KEMRI/ETHICS REVIEW COMMITTEE

KENYA MEDICAL RESEARCH INSTITUTE

P.O. Box 54840-00200, NAIROBI, Kenya
Tel (254) (020) 2722541, 2713349, 0722-205901, 0733-400003; Fax: (254) (020) 2720030
E-mail: director@kemri.org info@kemri.org Website:www.kemri.org

KEMRI/RES/7/3/1 **September 16, 2014**

TO: **MR. GODFREY BIGOGO,**
PRINCIPAL INVESTIGATOR

THROUGH: **DR. STEPHEN MUNGA,**
THE DIRECTOR, CGHR
KISUMU

Dear Sir,

RE: SSC PROTOCOL No. 2074: (*REQUEST FOR ANNUAL RENEWAL):* HUMAN GENETIC FACTORS AND CHILDHOOD MALARIA IN THE ASEMBO AREA OF WESTERN KENYA

Thank you for the Continuing Review Report for the period **11th October 2013 to 5th August 2014.**

This is to inform that during the 231st meeting of the KEMRI/ERC meeting held on the 19th of August 2014, the Committee **conducted the annual review and approved** the above referenced application for another year.

This approval is valid from today, **16th August 2014** through to **15th August 2015.** Please note that authorization to conduct this study will automatically expire on **15th August 2015** If you plan to continue with data collection or analysis beyond this date please submit an application for continuing approval to the **ERC** secretariat by **4th August 2015.**

You are required to submit any amendments to this protocol and other information pertinent to human participation in this study to the SSC and ERC for review prior to initiation.

Yours faithfully,

EAB

PROF. ELIZABETH BUKUSI,
ACTING SECRETARY,
KEMRI ETHICS REVIEW COMMITTEE

APPENDIX VI: Cross sectional survey form

7059277805

IBD Cross Sectional Survey Form (>12 year old)

Date of Interview(dd/mm/yyyy) ☐☐ / ☐☐ / ☐☐☐☐

Facility
- O Lwak
- O Mahaya
- O Ongielo

Demographics filled by: ☐☐

date16
Site
demoregcode

Part 1. DEMOGRAPHICS

1.0 Names of the Child:

First name ☐☐☐☐☐☐☐☐☐☐☐☐

Juok name ☐☐☐☐☐☐☐☐☐☐☐☐

Last name ☐☐☐☐☐☐☐☐☐☐☐☐

1.1a Child's DSS Permanent ID ☐☐☐-☐☐☐☐-☐☐☐-☐☐☐

fname
jname
lname
dssid

"Fill questions 1.1b and 1.1c only if the child is a new born and does not have DSS permanent ID."

1.1b Child's Temporary ID *(newborns)* ☐☐ ☐☐☐☐☐☐☐☐

1.1c Mother's DSS Permanent ID ☐☐☐-☐☐☐☐-☐☐☐-☐☐☐

Tempid
Motherid

1.2 Date of birth (dd/mm/yyyy) ☐☐ / ☐☐ / ☐☐☐☐

dob

1.3a How many years and months old is the child?
(If child < 1 month old, record # of days)

Years ☐☐ / Months ☐☐ / days ☐☐

1.0 Names of the household head:

First name ☐☐☐☐☐☐☐☐☐☐☐☐

Juok name ☐☐☐☐☐☐☐☐☐☐☐☐

Last name ☐☐☐☐☐☐☐☐☐☐☐☐

chlidyec
childmonth
childdays

1.4 For the head of household, what is the highest level of education completed?

- O None
- O Primary
- ☐☐☐ Years Completed
- O Higher
- O Secondary

1.5 What type of wall was used for the construction of this house? Observe

O Mud O Brick O Cement/Plastered O Cement/Painted O Other

1.6 What type of roof was used for the construction of this house?

O Grass thatch O Iron Sheet O Tiles O Other

1.7 Are the eaves open or closed? O Open O Closed O Partially Open

5773277806

1.8 What is the main material of the floor? **Observe**

Natural Floor	Rudimentary	Finished Floor
O Earth/sand O Dung	O Wood planks O Bambo/palm	O Parquet or polished O Wood vinyl or Asphalt strips O Ceramic tiles O Cement O Carpet

Other Specify ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

1.9 At any time in the past 12 months, has anyone sprayed the interior walls of your dwelling with chemical or liquid to kill or repel mosquitoes?(If No/Don't Know, skip next two questions)

O YES O NO O DON.T KNOW

1.9a How many months ago was the house sprayed? ☐☐☐ Months

1.9b Who sprayed the house?

Spray Agent

O PRIVATE COMPANY O HOUSEHOLD MEMBER O DON'T KNOW

O GOVERNMENT WORKER/PROGRAM O OTHER

Other Specify ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

2.0) Have any of the following been used in your house over the last week?

Mosquito Coil O yes O no Insecticide spray O yes O no Repellents O yes O no

2.1 Does your household have any mosquito nets that can be used while sleeping?

O YES O NO

Part II-Bednet Questionnaire:

2.2 : Ask respondent to show you the net in the house.

O Observed O Not Observed

2.2a Net description:

☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

2.3 How long ago did your household obtain the mosquito net?(Please input the specific number of months (0-36) if less than 3 years)

O Less than /equal to 3 years O More than 3 years ☐☐☐ Months

2.4 Observe or ask the brand of net:

O PermaNet O Supanet

O Olyset O Other ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

O SupaNet extra power O Don.t know brand

2.5 Since you got the mosquito net, was it ever soaked or dipped in a liquid to repel mosquitoes or bugs? N

O YES O NO O Not sure

2.6 How many months ago was this bed net treated with insecticide? ☐☐☐ Months

2.7. What did you treat the net with?

O PowerTab O KO Tab 123

O Don.t know O Other: ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

1994277807

Part II-Bednet Questionnaire:

2.8 Did anyone sleep under this net last night?

O YES O NO(skip to last question)

Part III:-Person Questionnaire.-

3.0 Client.s participation

O consent given O Refusal

O Dead O Not at home -

3..1 person answering

O SELF O PROXY

3.1b if proxy give reason

O person is under 15 years

O too ill/sick to talk

O other

3.2)Does the child usually live here?

O YES O NO

3.3 Did the child sleep here last night?

O YES O NO

3.4 What is the highest education level of the primary caretaker for the child?

O Primary O Higher

O Secondary O None ☐☐☐ Years Completed

3.5 Has the child been ill with a fever at any time in the last 2 weeks?

O YES O NO O NOT SURE

3.6 Did you seek advice or treatment for the fever from any source?

O YES O NO O DONT KNOW

3.7)Where did you seek advice or treatment? (Check all that apply)

O GOVT. HOSPITAL
O GOVT. HEALTH CENTER
O GOVT. DISPENSARY
O PVT./MISSION HOSPITAL/CLINIC
O PHARMACY/CHEMIST
O shop
O Freiend/realative
O private mobile clinic
O MOBILE CLINIC
O COMMUNITY HW
O OTHER PUBLIC
O PRIVATE DOCTOR/CLINIC
O Trad. Practitioner
O religious leader
O Other

3.8 Has the child taken any drugs in the last 2 weeks/did you take any medicine for the fever? (Check all that apply)

Acetaminophen/paracetamol/panadol/rofin	O yes			SP/FansidarOrodar/Metakelfin	O yes
Coartem	O yes	Quinine	O yes	Ibuprofen	O yes
Aspirin	O yes	Artemisinin	O yes	Septrin	O yes
Choloroquin	O yes	Malaratabs	O yes	Dont Know	O yes

Amodiaquine//Melaraquin O yes

Other Antimalarial ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Other Drug ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

9805277806

Part III:-Person Questionnaire.-

3.9 Has the child had a fever in the last 24 hours?

O YES O NO O DONT KNOW

Part IV-RDT/Anemia Questionnaire:-

4.0 Blood sample taken

O YES ,if YES skip to question 4.2
O No

4.1)Reason not taken

O REFUSED O NOT PRESENT O OTHER

4.2 RDT Result

O Positive
O Negative
O Invalid
O Not done
O Not applicable

4.3 Blood smear prepared

O YES
O No

4.4 EDTA sample collected?

O YES O NO O NOT APPLICABLE

4.4b Filter paper taken?

O YES O NO

4.5) Sample ID number(Record on blood slide and filter paper)

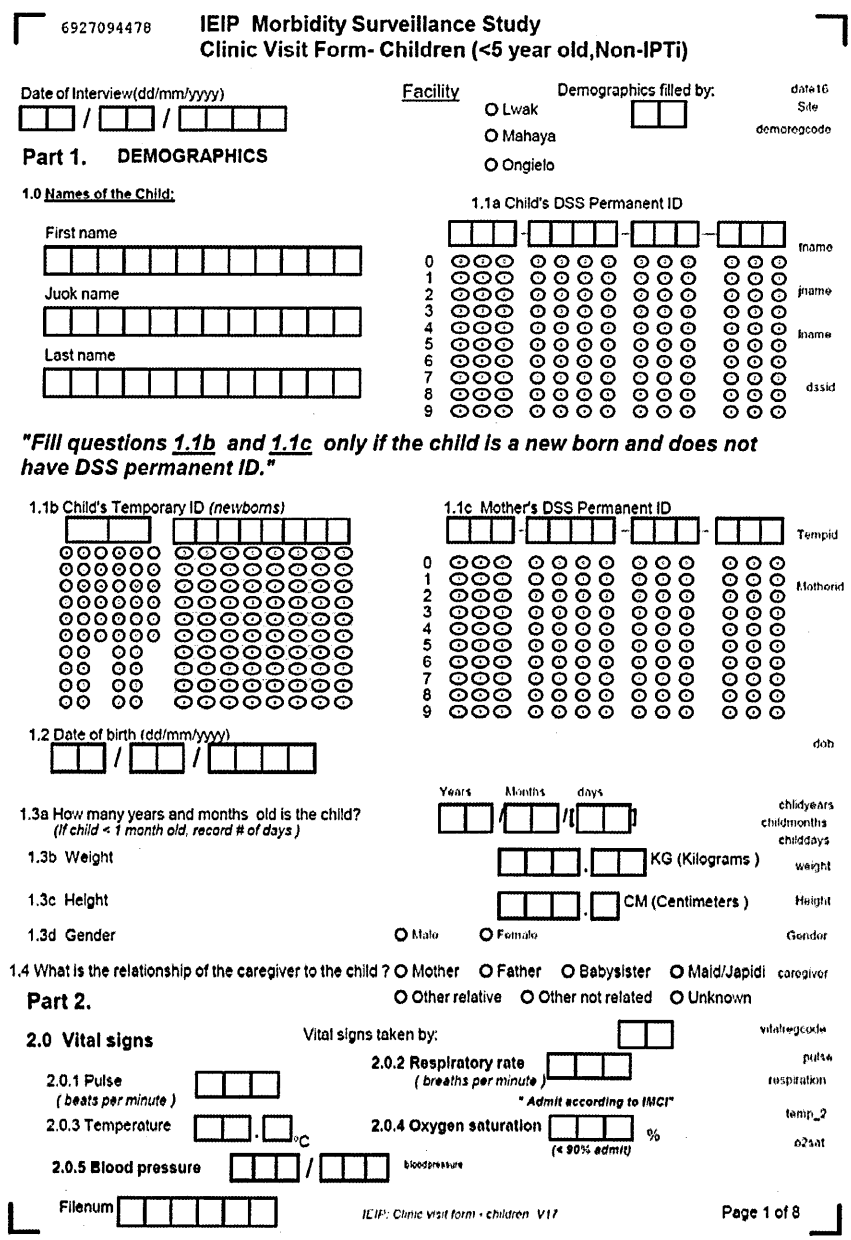
Place Barcode Here

#1 Finger pric/heel prick Label

4.6)Are there any other under 12's who live in this house that we have not tested?

O YES O NO

APPENDIX VII: IEIP child clinic visit form



6927094478

IEIP Morbidity Surveillance Study
Clinic Visit Form- Children (<5 year old,Non-IPTi)

Date of Interview(dd/mm/yyyy) ☐☐ / ☐☐ / ☐☐☐☐

Facility
O Lwak
O Mahaya
O Ongielo

Demographics filled by: ☐☐

date16
Site
demoregcode

Part 1. DEMOGRAPHICS

1.0 Names of the Child:

First name ☐☐☐☐☐☐☐☐☐☐☐☐

Juok name ☐☐☐☐☐☐☐☐☐☐☐☐

Last name ☐☐☐☐☐☐☐☐☐☐☐☐

1.1a Child's DSS Permanent ID ☐☐☐-☐☐☐☐-☐☐☐-☐☐☐

0 1 2 3 4 5 6 7 8 9 ○○○ ○○○ ○○○ ○○○ fname jname lname dssid

"Fill questions 1.1b and 1.1c only if the child is a new born and does not have DSS permanent ID."

1.1b Child's Temporary ID *(newborns)* ☐☐ ☐☐☐☐☐☐☐☐

1.1c Mother's DSS Permanent ID ☐☐☐-☐☐☐☐-☐☐☐-☐☐☐ Tempid

0 1 2 3 4 5 6 7 8 9 ○○○ ○○○ ○○○ ○○○ Mothord

1.2 Date of birth (dd/mm/yyyy) ☐☐ / ☐☐ / ☐☐☐☐ dob

1.3a How many years and months old is the child?
(If child < 1 month old, record # of days)
Years ☐☐ / Months ☐☐ / days ☐☐
childyears childmonths childdays

1.3b Weight ☐☐☐.☐☐ KG (Kilograms) weight

1.3c Height ☐☐☐.☐ CM (Centimeters) Height

1.3d Gender O Male O Female Gender

1.4 What is the relationship of the caregiver to the child ? O Mother O Father O Babysister O Maid/Japidi O Other relative O Other not related O Unknown caregiver

Part 2.

2.0 Vital signs

Vital signs taken by: ☐☐ vitalregcode

2.0.1 Pulse *(beats per minute)* ☐☐☐ pulse

2.0.2 Respiratory rate *(breaths per minute)* ☐☐☐ respiration

" Admit according to IMCI"

2.0.3 Temperature ☐☐.☐ °C temp_2

2.0.4 Oxygen saturation ☐☐☐ % *(< 90% admit)* o2sat

2.0.5 Blood pressure ☐☐☐ / ☐☐☐ bloodpressure

Filenum ☐☐☐☐☐☐

IEIP: Clinic visit form - children V17

Page 1 of 8

0713094475

Part 3 Treatment History

History and symptoms taken by: ☐☐ histregcode

3.1 What is the reason for today's visit ? *(Ango momiyo ikelo nyathi e ospital kawuono)* opdwhy

- O 1. need to consult since child is sick
- O 2. return for further treatment (new/worse illness)
- O 3. return after weekend consultation
- O 4. immunization (child is also ill)
- O 5. immunization (child is not ill)
- O 6. CCC Folow-up and Sick Visit
- O 10. Other: _____

3.1b How long have you had this symptoms? ☐☐☐ days

3.1.0 Presenting Symptoms *(Mark all that apply. Do not probe)* Symptoms

O Fever O Cough O Difficult breathing O Chest pain
O Sore throat O Diarrhea O Snezzing O Ear problem
O Headache O Burn O Yellow eyes/skin
O Other_____

3.1.1 Did the sick child visit anyone for health care before coming here? O Yes O No O Unknown visit

3.1.1.a If yes, which? *(Do not probe. Mark all that apply)* ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ipdhcb

Other Specify ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ OtherIpdhcb

(KEY)

1=Family Friend, 2= Traditional healer, 3=Nyamrerwa, 4=Bush Doctor,5 =Shop/Duka, 6=Health Facility, 7=Chemist, 8=Other.

3.2 Has the child taken any medications for this illness ? O Yes O No O Unknown medication

3.2a What was your child given?

				Duration					Duration		
SP	O yes	O no	O UKN	☐☐☐	Amodiaquine	O yes	O no	O UKN	☐☐☐	sp	aq
Chloroquine	O yes	O no	O UKN	☐☐☐	Analgesics	O yes	O no	O UKN	☐☐☐	cq	anal
Quinine	O yes	O no	O UKN	☐☐☐	Co-artem	O yes	O no	O UKN	☐☐☐	qu	co_artem
Septrin (CTX)	O yes	O no	O UKN	☐☐☐	Penicillin / amoxacillin	O yes	O no	O UKN	☐☐☐	ctx	pen
Other antibiotic	O yes	O no	O UKN	☐☐☐	Traditional Medicine	O yes	O no	O UKN	☐☐☐	antib	ipdmttm
Oral Rehydration Solution	O yes	O no	O UKN	☐☐☐							ORS
	(If yes, how many sachets ?)			☐☐							ors_sachet
Zinc	O yes	O no	O UKN	☐☐☐							Zinc
	(If yes, how many days ?)			☐☐							Zinc_days
Other:	O yes	O no	O UKN	☐☐☐							other

Medication name ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ ☐☐☐ treatname

Part 4. RISK FACTORS

4.1 Have you travelled outside Nyanza in the last 1 month? O Yes O No O unknown

If YES, to where

4.2 Has the patient traveled outside of Kenya in the last 1 month? O Yes O No O unknown travel

if yes mark all that apply:

O Sudan O Ethiopia O Tanzania O Rwanda O Somali ai_country

O Uganda O Burundi

Other 1: other1

Other 2: other2

Part 5. EXAMINATION (To be done by clinical officer or nurse)

Exams done by: examregcode

5.1.1 Is the child taking daily cotrimazole prophylaxis? O Yes O No O Unknown dailyctx

5.1.2 Is the child taking antiretroviral meds (ART)? O Yes O No O Unknown art

5.1.2b Did the child receive any vaccine in the last 7 days O Yes O No O Unknown **vaccine**

5.1 a Is this child able to drink or breastfeed at all? O Yes O No breastfeed_Co

5.1 b Does the child vomit everything with each feeding? O Yes O No Vomitany_Co

5.2 Does the child have difficult breathing ? O Yes O No diffbreathsign

5.3 Does the child have chest pain when breathing ? O Yes O No O Unknown Chestpain

5.4 Has the child had hot body or fever with this illness? O Yes O No O Unknown fever

(del maore / maliet)

5.4 a If yes, how many days ? days feverdays

5.5 Does the child have cough ? *(ahonda)* O Yes O No O Unknown cough

5.5 a. If yes, how long has the child had cough ? days coughdays

5.5 b. If yes, coughing blood ? (fuolo remo) O Yes O No O Unknown coughblood

5.6 Does the child have diarrhea ? *(diep)* O Yes O No O Unknown diarrhea

5.6 a. If yes, how long has the child had diarrhea ? days diarrheadays

5.6 a 1 How many loose stool did the child have in the last 24 Hrs? *Hrs* LooseStool

5.6b Is the stool watery? O Yes O No O Unknown Waterystool

5.6 b. Is there blood in the stool ? O Yes O No O Unknown stool

5.6 b. Is there blood in the stool ? *(diep mar remo)* O Yes O No O Unknown stool

5.6 c. Is there mucus in the stool ? *(diep ma kirendarenda)* O Yes O No O unknown diarrmuc

5.7 Does the child have a runny nose ? *(um mamol)* O Yes O No O Unknown runnynose

5.8 Does the child have sneezing ? *(jir)* O Yes O No O Unknown sneezing

5.9 Does the child have an ear problem ? *(it machendo)* O Yes O No O Unknown earproblem

6243094479

5.9 a. If yes, how many days ?	☐☐☐ days	eardays
5.10 Does the child have sore throat ? *(duol maremo)*	O Yes O No O Unknown	sorethroat
5.11 Has the child had dark urine ? *(lach rabuor)*	O Yes O No O unknown	darkurine
511 b. Has the child had yellow eyes with this illness ?	O Yes O No O unknown	jaundicehx
5.11 c. If yes, how many days ?	☐☐☐ days	jaundicedays
5.12 Has the child had pale stool ? *(oko ma marmar)*	O Yes O No O unknown	palestoo
5.13 Is the child lethargic ?	O Yes O No	lethargic
5.13 a) Is this child unconscious?	O Yes O No	ipdment
5.13 b AVPU Scale	O A O V O P O U	avpu
5.14 a Is the child convulsing now ?	O Yes O No	convulsion
5.14 b Has the child had convulsions with this illness?	O Yes O No	convulsion_Co
5.15 Is the child not able to drink or drink poorly ?	O Yes O No	poordrink
5.16 Is the child excessively breastfed ?	O Yes O No	BreastFeed
.050.16.1 Has the child vomited with this illness ?	O Yes O No	Vomit
5.17a Does the child have fast breathing? O Yes O No	***If <2months and >59 bpm*** --- ***severe disease*** ***If 2-12months and >49 bpm*** } --- ***pneumonia*** ***If >12 months and >39 bpm***	breaths
5.17 b Chest indrawing?	O Yes O No	indrawing
5.17 c Nasal flaring?	O Yes O No	flaring
5.17 d Stridor when calm?	O Yes O No	stridor
5.17 e Wheezing?	O Yes O No	wheezing
5.17 f Crackles/rales?	O Yes O No	crackles
5.17 g Grunting?	O Yes O No	grunting
5.18 Is the child restless or irritable ?	O Yes O No	restless
5.19 Does the child have sunken eyes ?	O Yes O No	sunkeneyes
5.20 Is the child drinking eagerly, thirsty ?	O Yes O No	thisty
5.20 b Does your child have tears when he/she cries ?	O Yes O No O N/A	tears
5.21 After pinching the skin, how does it go back ?	O Very slowly O Slowly O Normally	pinch
5.2 Does the child have stiff neck ?	O Yes O No	stiffneck
5.23 Does the child have bulging fontanelle ?	O Yes O No	fontonelle
5.23b. Does the child a sunken fontanelle ?	O Yes O No	fontonelsunk
5.24 Does the child have rash ?	O Yes O No	rash
5.24a. describe the rash ?	O Heat O Diaper O Maculopapular O Vesicular O Petechial O Abscess O Infected (purulent skin lesion)	diaperrash
5.24.b other kind of Rash	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐	diaperrashdes
5.25 Does the child have red eyes ?	O Yes O No	redeyes
5.26 Does the child have mouth ulcers ?	O Yes O No	mouthulcers

3342094476

5.27 Does the child have pus draining from the eyes ? O Yes O No puseyes

5.27b. Does the child have pus draining from the ears ? O Yes O No pusears

5.27c. Was ear exam done ? O Normal O Not normal O Not done earexam

5.27d. If not normal O Discharge O Inflamation O Any other earxam_notnomal

5.28 Is there visible severe wasting ? O Yes O No wasting

5.29 Is there oedema of both feet ? O Yes O No oedema

5.30 Does the child have jaundice ? O Yes O No jaundice

5.31 Are there enlarged lymph nodes at 2 or 3 of the following sites, Neck, Axillae,Groin. ? O Yes O No lymphnodes

5.32 Oral thrush ? O Yes O No oralthrush

5.33 Any other sign of anemia (paleness of palms, nailbed, conjunctiva) ? O Yes O No pale

5.34 Other symptoms 1. othersymp1

5.35 Other symptoms 2. othersymp2

Part 6. SAMPLES

6.0 Case defination met: O None O SARI O ILI O Fever O Diarrhea O Admission O Jaundice CaseDefintion

6.0 a Patient eligible for : O NP/OP O Blood Culture O Serum O Stool O Urine O X-rays SampEligibility

6.0 b Was chest X-ray ordered? O Yes O No *If No, why?* O Patient is pregnant O Done in last 3months XRayOrdered

O Others (specify) XrayNo XrayNoOther

6.0 b Nurse Code NurseCode

hb_16

6.1 Haemoglobin . g/dl

6.2 Blood slide done ?
O Yes O Refused O Not done

6.3b Giemsa stain BS results?
O Positive O Negative

6.3c parasite species? O Falciparum O Malariae O Ovale O Vivax

6.3d parasite count per 200 wbc paracount

6.4 Nasopharyngeal swab O Yes O Refused O Not done npswab
Not done why?

6.4a Oropharyngeal swab O Yes O Refused O Not done orswab
Not done why?

6.5 a. Stool sample taken O Yes O Refused O Not done bloodslide
Not done why? stoolsamp

6.6 Serum sample taken ? O Yes O Refused O Not done sera
Not done why? giemsa

6.7 Blood culture done ? O Yes O Refused O Not done bloodcx

malspec

X- RAY RESULT

X-ray done by:

6.3 e) Is the filim quality O Adequate O Suboptimal O Unreadable Filim_quality

6.3 f) Does the filim contain significant pathology ? O Yes O No O Not Done Pathology

6.3 g) Primary end Point consolidation ? Right O Yes O No Right_Consolidation

Left O Yes O No Left_Consolidation

6.3 h) Other consolidation / infiltrate ? Right O Yes O No Right_Infiltrat

Left O Yes O No Left_Infiltrat

6.3 i) Pleural fluid ? Right O Yes O No Pleural_Right

Left O Yes O No Pleural_Left

6.3 j) Other abnormalities ? Other_Abnormalities

Filenum

2235094476

Part 7: Diagnosis

7.01 What do you diagnose the child with?

Malaria	O yes	Dysentery	O yes	svmalaria svdysentery
Pneumonia/Lower respiratory tract infection	O yes	Intestinal worms	O yes	svpneumonia svworms
Upper Respiratory Tract Infection(URTI)	O yes	Anemia	O yes	svurti svanemia
Wheezing/bronchospasm	O yes	Malnutrition	O yes	svbronchospasm svmalnutrition
Otitis media	O yes	Oral candidiasis	O yes	svotitis svcandidiasis
Conjunctivitis	O yes	Rash/skin problem	O yes	svdiarrhoea3
Meningitis	O yes	Scabies	O yes	svconjuctiva
Diarrhoea/ Gastroenteritis	O yes	Burn	O yes	svrashd svmeningitis svscabies
Dehydration	O yes	Wound/injury	O yes	svburn
Pharyngitis/tonsillitis	O yes	Amoebiasis	O yes	svdehydration svwound svPharyngitis
Viral syndrome	O yes	Pulmonary TB	O yes	svamoebiasis
Convulsions	O yes	Influenza	O yes	svViralsyn svptb ipdacon
Measeles	O yes	Urinary Track Infection	O yes	svflu ipdams1
Other1				otherdx1

7.02 Was there any error in question 7.01? O yes ● no — anyerror6_1

11.5 Did the patient have an HIV test as part of DCT/PITC during the visit? O Yes O No O Don't Know — hivtest

11.5a) If No, was it because: O Illness not suspected to be HIV related O Already known to be HIV — hivtestno

O Patient refused

O Other — hivtestother

11.5b) If yes,(*Only answer if consent obtained*) what was the result of DCT/PITC — DCTresult

O Positive O Negative O Interminate

11.5c) Did the child have blood for PCR taken ? O Yes O No O N/A — pcrtaken

11.5d)Was PCR O Positive O Negative O Unknown — PCRresult

Part 8: MEDICATION

Antimalarials

8.1 Did you or will you give antimalarial treatment today? O yes O no ***(If 'no', go to 8.2.)*** — antimalaria_16

8.1a What was given?

SP	O yes	Qty	days	Amodiaquine	O yes	Qty	days
Coartem	O yes	Qty	days	Quinine	O yes	Qty	days

msp
maq
mco_artem
mqu

8.2 Did you or will you give antibiotics treatment today? O yes O no ***(If 'no', go to 8.3.)*** — antibiotics_16

8.2a What was given?

Septrin	O yes	Qty	days	Amoxacillin/ ampicillin	O yes	Qty		Penicillin	O yes	Qty	
Gentamicin	O yes	Qty	days	Tetracycline	O yes	Qty		Metronidazole	O yes	Qty	

mctx
mamox
mpen
mgent
mtetra
mmet

Filenum

7960094476

Ciprofloxacin O yes Qty days Nalidixic Acid O yes Qty Antihelminthic O yes Qty mchlory mhal mantih

Erythomycin O yes Qty days Cloxacillin O yes Qty Ceftriaxone O yes Qty merythro mclo

Chloramphe O yes Qty days Doxycyclin O yes Qty mceft mchloro mdoxy

Other Antibiotic antibiother_16

8.3 Did you or will you give other treatment today? O yes O no *(If 'no', go to 9.)* mother_16

8.3a What was given?

Analgesic O yes Salbutamol/ventolin O yes Piriton O yes MVI O yes manal

Oral Rehydration solution O yes Valium O yes Expectorant O yes Zinc O yes mors msalbutamol mval mmvi mpiriton mexpect mzinc

Other Medicin1 othermed1

Other Medicin2 othermed2

Part 9: DISPOSITION (Admit for any symptoms or signs in bold.)

9.1 What is the disposition of the child ?

O Home O Admit Lwak O Refer Bondo District Hospital O Refused admission disposition

O Refer PGH

O Other *(specify)* dispother

9.2 Follow up appointment O None O Hours hrs O Days days O Weeks wks O Months mnths O Years yrs

Part 10: HOSPITAL COURSE

10.1 Did you give antimalarial medicines during the hospital course ? O yes O no *(If 'no', go to 10.2.)* hantimalaria_16

10.1a What was given? hmsp

Sp	O yes	Amodiaquine	O yes		hmaq
Coartem	O yes	Quinine	O yes		hmco_arte hmq

10.2 Did you give antibiotics medicines during the hospital course ? O yes O no *(If 'no', go to 10.3.)* hantibiotics_16

10.2a What was given?

Septrin	O yes	Amoxacillin/ampicillin	O yes	Penicillin	O yes	hmamox hmct hmgen hmpen
Gentamicin	O yes	Tetracycline	O yes	Metronidazole	O yes	hmmet hmtetra
Ciprofloxacin	O yes	Nalidixic Acid	O yes	Antihelminthic	O yes	hmcipro hmnal hmantih
Ethomycin	O yes	Cloxacillin	O yes	Ceftriaxone	O yes	hmerythro
Chloramphenicol	O yes	Doxycyclin	O yes			hmclox hmceft hmchloro

10.2b Did you give intravenous fluids? O yes O no 10.2b1 How many bottles given hmdoxy ivfluids ivbottles

10.3 Did you give other medicines during the hospital course ? O yes O no

10.3a What was given? *If 'no', go to 11.* hmanal hmother_16

Salbutamol/ventolin	O yes	**Zinc**	O yes	**Analgestic**	O yes	hmsalbutamol
Oral Rehydration solution	O yes	**Valium**	O yes	**Multivitamin**	O yes	hmmultivit hmors hmval

Other Medicin1 hmzinc hothermed1

Other Medicin2 hothermed2

Part 11: DISCHARGE

11.1 Date of discharge/death? (dd/mm/yyyy) / /

11.2 Outcome of hospital course ? hospddate

O Discharge home without sequelae O Absconded O Died

O Discharge home with sequelae O Transferred hospoutcome

Filenum

8886094474

Filenum

11.3 What is the discharge diagnosis ?

Malaria	O yes	Dysentery	O yes	hdalaria hddysentery
Pneumonia/Lower respiratory tract infection	O yes	Intestinal worms	O yes	hdpneumonia
Upper Respiratory Tract Infection(URTI)	O yes	Anemia	O yes	hdworms hdurti hdanemia
Wheezing/bronchospasm	O yes	Malnutrition	O yes	hdbronchospasn
Otitis media	O yes	Oral candidiasis	O yes	hdmalnutrition hdotitis
Conjunctivitis	O yes	Rash/skin problem	O yes	hdcandidiasis
Meningitis	O yes	Scabies	O yes	hdconjuctiva hdrashd
Diarrhoea/ Gastroenteritis	O yes	Burn	O yes	hdmeningitis hdscabies
Dehydration	O yes	Wound/injury	O yes	hdburn hddehydration
Convulsions	O yes	Pulmonary TB	O yes	hdwound hdptb
Measles	O yes	Influenza	O yes	ipddcon hdflu
UTI	O yes			ipddms1

svotherdx1

Other1

11.4 Was there any error in question 10.3 ? O yes ● no

Specimen Label Intake

12.1 Has specimen been taken ? O yes O no specimenlabel

Time of blood draw Time_BloodDraw

Place White Portion Here

#1 Nasopharyngeal Swab /Oropharyngeal Swab Label

NP/OP Time

Place White Portion Here

#2 Blood, Serum Label

Serum Time

#5 Blood, Culture Label

Culture Time

Place White Portion Here

Place White Portion Here

#3 Stool Sample Label

Stool Time

6 Blood culture Label Myco/ F-Lytic

Myco/FLytic Time

Place White Portion Here

Place White Portion Here

4Urine Sample Label

Urine Time

7 Blood Clot

Clot Time

Place White Portion Here

Filenum